

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 01:36:54 ; Search time 175 Seconds  
(without alignments)  
2651.561 Million cell updates/sec

Title: US-10-619-992-2

Perfect score: 6215

Sequence: 1 MPVQAQWTEFLSCPICYNE.....MSDKNDFLKPVANGRMVNS 1191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1745140 seqs, 389608008 residues

Total number of hits satisfying chosen parameters: 1745140

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10F\_PUBCOMB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| 1          | 6215  | 100.0       | 1191   | 9  | US-09-921-099-2      |
| 2          | 6207  | 99.9        | 1191   | 9  | US-09-921-099-4      |
| 3          | 611   | 9.8         | 522    | 15 | US-10-104-047-3665   |
| 4          | 490   | 7.9         | 95     | 9  | US-09-867-550-1346   |
| 5          | 311   | 5.0         | 677    | 9  | US-09-864-761-48290  |
| 6          | 207   | 3.3         | 2377   | 18 | US-10-840-512-196    |
| 7          | 203   | 3.3         | 1267   | 15 | US-10-092-900A-128   |
| 8          | 200.5 | 3.2         | 728    | 17 | US-10-732-923-6679   |
| 9          | 198.5 | 3.2         | 3586   | 15 | US-10-334-143-77     |
| 10         | 196.5 | 3.2         | 1248   | 16 | US-10-437-963-126940 |
| 11         | 195.5 | 3.1         | 123    | 9  | US-09-864-761-42017  |

ALIGNMENTS

RESULT 1  
US-09-921-099-2  
; Sequence 2, Application US/09921099  
; Patent No. US20020107372A1  
; GENERAL INFORMATION:  
; APPLICANT: Hefeneider, Steven  
; APPLICANT: Merkins, Louise  
; APPLICANT: Bennett, Robert  
; APPLICANT: Seiss, Donald  
; TITLE OF INVENTION: Mammalian DNA Binding Membrane-Associated Protein-encoding Gene  
; FILE REFERENCE: 00-617-A  
; CURRENT APPLICATION NUMBER: US/09/921,099  
; CURRENT FILING DATE: 2001-08-01  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 2  
; LENGTH: 1191  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-921-099-2

Query Match 100.0%; Score 6215; DB 9; Length 1191;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MPVQAQWTEFLSCPICYNEFDENVHKPISLGSHTVCKTCLNKLHKKACPFDDTAINTD 60  
Db 1 MPVQAQWTEFLSCPICYNEFDENVHKPISLGSHTVCKTCLNKLHKKACPFDDTAINTD 60  
QY 61 IDVLPVNFALLQLVGAQVDPHQSIKLSNLGENKHVEYAKKCVEDLALYLPKLSGGKGVAS 120  
Db 61 IDVLPVNFALLQLVGAQVDPHQSIKLSNLGENKHVEYAKKCVEDLALYLPKLSGGKGVAS 120  
QY 121 LNSALSRRPMQRKLVTLVNLQVLEEGRVAMRAARSLGERTVTTELQHNFPQULSANL 180

Db 121 LNQSALSRPMQKLVTLVNCQVEBGRVRAARAALSGERTVTELLIHOHQNPQQLSANTL 180  
QY 181 WAAVARGCQFLGPANQBEALKLVLALBDGSGALSRKVLVLFVQVQLEPRFPQASKTSG 240  
Db 181 WAAVARGCQFLGPANQBEALKLVLALBDGSGALSRKVLVLFVQVQLEPRFPQASKTSG 240  
QY 241 HVVQLLYRASCFTKTRDEDSIMOLKEBFRSYEARLREHDAQIVHIAEAGLRISPEQM 300  
Db 241 HVVQLLYRASCFTKTRDEDSIMOLKEBFRSYEARLREHDAQIVHIAEAGLRISPEQM 300  
QY 301 SSLLYGDLAHKSHMQSIIDKLQSPSPFAKSVQELTIIVLQRTGDPANLNRLRPHLELLANI 360  
Db 301 SSLLYGDLAHKSHMQSIIDKLQSPSPFAKSVQELTIIVLQRTGDPANLNRLRPHLELLANI 360  
QY 361 DNPDAVSTPWEQLENAMVAVKTVVHGLVDFIQTNSYRKGHETPQOPNSKYKTSMCRDLR 420  
Db 361 DNPDAVSTPWEQLENAMVAVKTVVHGLVDFIQTNSYRKGHETPQOPNSKYKTSMCRDLR 420  
QY 421 QGGCPRGTNCTFAHSHQSELEKYLRNKKINATVTRFPLLNKVGNNVTVTAGNVSIV 480  
Db 421 QGGCPRGTNCTFAHSHQSELEKYLRNKKINATVTRFPLLNKVGNNVTVTAGNVSIV 480  
QY 481 GSTETTGKIVPSTNGISNAENSVSOLISRSTDSLRALETVKVKGVGANGQNAAGPSAD 540  
Db 481 GSTETTGKIVPSTNGISNAENSVSOLISRSTDSLRALETVKVKGVGANGQNAAGPSAD 540  
QY 541 SVTENKIGSPKTPVSNVAATAGPSNVGTSLNSVPQKSSPFLTRVPVYPHSENIQYFQ 600  
Db 541 SVTENKIGSPKTPVSNVAATAGPSNVGTSLNSVPQKSSPFLTRVPVYPHSENIQYFQ 600  
QY 601 DRTQIPFEVPOYPTGYPPPTVPAGVPCVPRFVRNNVPSESLPPASMPYADHYST 660  
Db 601 DRTQIPFEVPOYPTGYPPPTVPAGVPCVPRFVRNNVPSESLPPASMPYADHYST 660  
QY 661 FSPDRMNSSPYQPPPPQYGPVPPVPSGMYPVVDSSRIWRPMPYQORDDIIRSNLPPM 720  
Db 661 FSPDRMNSSPYQPPPPQYGPVPPVPSGMYPVVDSSRIWRPMPYQORDDIIRSNLPPM 720  
QY 721 DVMHSSVYQTSLRERYNSLDGYYSVACQPPSPRTTVPVLPREPCGHLKTSCEQIRRKPD 780  
Db 721 DVMHSSVYQTSLRERYNSLDGYYSVACQPPSPRTTVPVLPREPCGHLKTSCEQIRRKPD 780  
QY 781 QNAQYHTQKAPLVSSSLPVATQSPPTPSPLFSVDREADESESUGTKFEEDHLSHYSPWS 840  
Db 781 QNAQYHTQKAPLVSSSLPVATQSPPTPSPLFSVDREADESESUGTKFEEDHLSHYSPWS 840  
QY 841 CGTIGSCINAIIDSEPKDVIANSNAVLMDLSDGDVKKRVHLFETQRTKEEDPIIIPSDGP 900  
Db 841 CGTIGSCINAIIDSEPKDVIANSNAVLMDLSDGDVKKRVHLFETQRTKEEDPIIIPSDGP 900  
QY 901 IISKWGAISRSRTGYHTTDPVQATASQGSATKPIISVSDYVPVNAVDSRWSSYNGEATS 960  
Db 901 IISKWGAISRSRTGYHTTDPVQATASQGSATKPIISVSDYVPVNAVDSRWSSYNGEATS 960  
QY 961 SAHYVERDRPIVTDLSGHRKHSSTGDLISLEQQAQKSNLILLOREANALAMOOKNSLDE 1020  
Db 961 SAHYVERDRPIVTDLSGHRKHSSTGDLISLEQQAQKSNLILLOREANALAMOOKNSLDE 1020  
QY 1021 GRHLTLNLISKELELNGELQSDYTDATDTPDRDIELELSALDTPDQSEPIEEL 1080  
Db 1021 GRHLTLNLISKELELNGELQSDYTDATDTPDRDIELELSALDTPDQSEPIEEL 1080  
QY 1081 DIQLGISNQDQLLNGMAVENGHVPQHQKEPPKQKQSLGEDHVILIEQKTLIPVTSCEF 1140  
Db 1081 DIQLGISNQDQLLNGMAVENGHVPQHQKEPPKQKQSLGEDHVILIEQKTLIPVTSCEF 1140  
QY 1141 SQPLPVSISNASCLPITTSVAGNLILKTHVWSEDKNDFLKPVANGKMYNS 1191  
Db 1141 SQPLPVSISNASCLPITTSVAGNLILKTHVWSEDKNDFLKPVANGKMYNS 1191

US-09-921-099-4  
; Sequence 4, Application US/09921099  
; Patent No. US20020107372A1  
; GENERAL INFORMATION:  
; APPLICANT: Hefeneider, Steven  
; APPLICANT: Merkins, Louise  
; APPLICANT: Bennett, Robert  
; APPLICANT: Seigs, Donald  
; TITLE OF INVENTION: Mammalian DNA Binding Membrane-Associated Protein-encoding Gene  
; FILE REFERENCE: 00-617-A  
; CURRENT APPLICATION NUMBER: US/09/921,099  
; CURRENT FILING DATE: 2001-08-01  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 1191  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-921-099-4

Query Match 99.9%; Score 6207; DB 9; Length 1191;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MPVQAAQWTEFLSCPICYNEDFNHVKPISLGCSTHTVCKTCLNKLHRRKACPDQTAINTD 60  
Db 1 MAVQAAQWTEFLSCPICYNEDFNHVKPISLGCSTHTVCKTCLNKLHRRKACPDQTAINTD 60  
QY 61 IDVLVFNFALLQLVGAQVDPHQSIKLSNLGENKHVEYAKKVEDLALYLKPLSGGKGVAS 120  
Db 61 IDVLVFNFALLQLVGAQVDPHQSIKLSNLGENKHVEYAKKVEDLALYLKPLSGGKGVAS 120  
QY 121 LNQSALSRPMQKLVTLVNCQVEBGRVRAARAALSGERTVTELLIHOHQNPQQLSANTL 180  
Db 121 LNQSALSRPMQKLVTLVNCQVEBGRVRAARAALSGERTVTELLIHOHQNPQQLSANTL 180  
QY 181 WAAVARGCQFLGPANQBEALKLVLALBDGSGALSRKVLVLFVQVQLEPRFPQASKTSG 240  
Db 181 WAAVARGCQFLGPANQBEALKLVLALBDGSGALSRKVLVLFVQVQLEPRFPQASKTSG 240  
QY 241 HVVQLLYRASCFTKTRDEDSIMOLKEBFRSYEARLREHDAQIVHIAEAGLRISPEQM 300  
Db 241 HVVQLLYRASCFTKTRDEDSIMOLKEBFRSYEARLREHDAQIVHIAEAGLRISPEQM 300  
QY 301 SSLLYGDLAHKSHMQSIIDKLQSPSPFAKSVQELTIIVLQRTGDPANLNRLRPHLELLANI 360  
Db 301 SSLLYGDLAHKSHMQSIIDKLQSPSPFAKSVQELTIIVLQRTGDPANLNRLRPHLELLANI 360  
QY 361 DNPDAVSTPWEQLENAMVAVKTVVHGLVDFIQTNSYRKGHETPQOPNSKYKTSMCRDLR 420  
Db 361 DNPDAVSTPWEQLENAMVAVKTVVHGLVDFIQTNSYRKGHETPQOPNSKYKTSMCRDLR 420  
QY 421 QGGCPRGTNCTFAHSHQSELEKYLRNKKINATVTRFPLLNKVGNNVTVTAGNVSIV 480  
Db 421 QGGCPRGTNCTFAHSHQSELEKYLRNKKINATVTRFPLLNKVGNNVTVTAGNVSIV 480  
QY 481 GSTETTGKIVPSTNGISNAENSVSOLISRSTDSLRALETVKVKGVGANGQNAAGPSAD 540  
Db 481 GSTETTGKIVPSTNGISNAENSVSOLISRSTDSLRALETVKVKGVGANGQNAAGPSAD 540  
QY 541 SVTENKIGSPKTPVSNVAATAGPSNVGTSLNSVPQKSSPFLTRVPVYPHSENIQYFQ 600  
Db 541 SVTENKIGSPKTPVSNVAATAGPSNVGTSLNSVPQKSSPFLTRVPVYPHSENIQYFQ 600  
QY 601 DRTQIPFEVPOYPTGYPPPTVPAGVPCVPRFVRNNVPSESLPPASMPYADHYST 660  
Db 601 DRTQIPFEVPOYPTGYPPPTVPAGVPCVPRFVRNNVPSESLPPASMPYADHYST 660  
QY 661 FSPDRMNSSPYQPPPPQYGPVPPVPSGMYPVVDSSRIWRPMPYQORDDIIRSNLPPM 720  
Db 661 FSPDRMNSSPYQPPPPQYGPVPPVPSGMYPVVDSSRIWRPMPYQORDDIIRSNLPPM 720  
QY 721 DVMHSSVYQTSLRERYNSLDGYYSVACQPPSPRTTVPVLPREPCGHLKTSCEQIRRKPD 780

Db 721 DVMSVYQTSURYNLDGYGVACQPPSPRTTVPFRPCGHLKTSCEQIRRPD 780  
Qy 781 QWAOYHTOKAPLVSSTLPVATOSPPTPPSPFLSVDFRADPSESVSCTKFEEDHLSHYSPWS 840  
Db 781 QWAOYHTOKAPLVSSTLPVATOSPPTPPSPFLSVDFRADPSESVSCTKFEEDHLSHYSPWS 840  
Qy 841 CQTIGSCINAIDSEPKDVIANAVLMDLSDGDKVRRVHLFTQRTKEEDPIIFPSDGP 900  
Db 841 CQTIGSCINAIDSEPKDVIANAVLMDLSDGDKVRRVHLFTQRTKEEDPIIFPSDGP 900  
Qy 901 IISKWGAISRSRTGHTTDPVQATASQGSATKPIISVDYVYVNAVDSRWSSSYNGEATS 960  
Db 901 IISKWGAISRSRTGHTTDPVQATASQGSATKPIISVDYVYVNAVDSRWSSSYNGEATS 960  
Qy 961 SAHYVERDRFIVTDLSGHKKHSTGDLLELQAKSNLLQREANALAMQKWNLSDE 1020  
Db 961 SAHYVERDRFIVTDLSGHKKHSTGDLLELQAKSNLLQREANALAMQKWNLSDE 1020  
Qy 1021 GRHLTLNLSKEIERNGLQSDYTEDATDTKPRDRIEELSALDTDEPDGQSEPIEIL 1080  
Db 1021 GRHLTLNLSKEIERNGLQSDYTEDATDTKPRDRIEELSALDTDEPDGQSEPIEIL 1080  
Qy 1081 DIQLGISSQNDQLLGMVANGHPVQHQKEPPKQKQSLGSDHVLBEQKTLIPVTSQF 1140  
Db 1081 DIQLGISSQNDQLLGMVANGHPVQHQKEPPKQKQSLGSDHVLBEQKTLIPVTSQF 1140  
Qy 1141 SQPLPVSISNASCLPITTSVSAGNLLKTHVMSDKNDFLKPVANGKWN 1191  
Db 1141 SQPLPVSISNASCLPITTSVSAGNLLKTHVMSDKNDFLKPVANGKWN 1191  
RESULT 3  
US-10-104-047-3665  
; Sequence 3665, Application US/10104047  
; Publication No. US20030236392A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: NO. US20030236392A1 full length cDNA  
; FILE REFERENCE: H1-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; CURRENT FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 3665  
; LENGTH: 522  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-104-047-3665  
Query Match 9.8%; Score 611; DB 15; Length 522;  
Best Local Similarity 34.1%; Pred. No. 2e-33;  
Matches 187; Conservative 76; Mismatches 187; Indels 98; Gaps 27;  
Qy 618 YTPPTTVAGVACVPRFRVRSNNVSESLPPASMEYADHYSTFSDRDMNSPIQPPPP 677  
Db 2 YTPPPP-----QCVSFRV-----PPPSAPEAPPYLDHYPPYL-OERVNSQYGTQPO 49  
Qy 678 QPYGVPVPPSGMYAPVYDSRRIRWPPMYORDDIIRSNLPPMDVMHSSV--YQTSRER 735  
Db 50 Q-----YPII-----YPSHYDGRVYPAVSYTREIFRESPI-PIEIPPAVPSVPSRER 100  
Qy 736 YNSLDGYGVACQF-----PS-----EPRTTVPLPREPCGHLKTSCEQIRRPDQWAOYHT 787  
Db 101 YQIESYYPVAPHTQIRPSYLRPPVSRLLPPPPQ-----HPSLDELHRRRKEIMAQLE 156  
Qy 788 QKAPLVSSTLPVATQSTPPSPFLSVDFRADPSESVSCTKFEEDHLSHYSPWSCTIGSC 847  
Db 157 RK---VISPPPPFA-PSPTLP-PTFHPPEEFLDEDLKVG-KYKNDYSQYSPWNSCDTIGS 210  
Qy 848 INADSEPKDVIANAVLMDLSDGDKVRRVHLFTQRTKE--EDPIIFPSDGPISKW 905

Db 211 IGTDAKPKOVAAAGSVEMMNVESKGM--RQRLDLQRAAETSDDDLIPGDRPTVSRF 268  
Qy 906 GAISRSRTGHTTDPVQATASQGSATKPIISVDYVYVNAVDSRW--SSYG--NEATSS 961  
Db 269 GAISRTKTIYQAGPQMAPOGAPTAKSINISDYPY--GTHGGMGASPYSPHQNIPSQ 326  
Qy 962 AHYVERDRFIVTDLSGHKK--HSSTGDLLELQ-----AKSNLLQOR 1004  
Db 327 GHFSERERISMSEVASHGKPLPSAEREQLRLELQLNHQISQOTQLRGPEAVSNRLVQR 386  
Qy 1005 EANALAMQ-----KNSLDSGRHLTLNL--LSKEIERNGLQSDYTEDATDTKPD 1054  
Db 387 EANTLAGOSQPPPPPPKWPCKMISSEQLSLELHQVEREIGKTRLSME--NCSLDMK-- 443  
Qy 1055 RDIELELSALDTDE--PDGQSEPIEILDIQLGISSQNDQLLGMVANGHPVQHQKEP 1112  
Db 444 -----SKLNTSKQAENGQPEPQNKV-----PAEDLTITFSDVPNGSALTQENISL 488  
Qy 1113 PKQKQSL 1120  
Db 489 LSNKTSSL 496  
RESULT 4  
US-09-867-550-1346  
; Sequence 1346, Application US/09867550  
; Patent No. US20020082206A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Mehraban, Fuad,  
; APPLICANT: Conley, Pamela  
; APPLICANT: Law, Debbie  
; APPLICANT: Topper, James  
; TITLE OF INVENTION: NO. US20020082206A1el Polynucleotides from Atherogenic Cells and  
; TITLE OF INVENTION: Thereby  
; FILE REFERENCE: 21402-013 (Cura-313)  
; CURRENT APPLICATION NUMBER: US/09/867,550  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: USSN 60/208,427  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 2125  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1346  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-867-550-1346  
Query Match 7.9%; Score 490; DB 9; Length 95;  
Best Local Similarity 100.0%; Pred. No. 3.5e-26;  
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1097 MAVENGHPVQHQKEPPKQKQSLGSDHVLBEQKTLIPVTSFQSPPLPVSISNASCLPI 1156  
Db 1 MAVENGHPVQHQKEPPKQKQSLGSDHVLBEQKTLIPVTSFQSPPLPVSISNASCLPI 60  
Qy 1157 TTSVAGNLLKTHVMSDKNDFLKPVANGKWN 1191  
Db 61 TTSVAGNLLKTHVMSDKNDFLKPVANGKWN 95  
RESULT 5  
US-09-864-761-48290  
; Sequence 48290, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

```
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48290
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL136170.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
; OTHER INFORMATION: SWISSPROT HIT: Q01970, EVALUE 2.60e+00
; OTHER INFORMATION: EST_HUMAN HIT: AW962574.1, EVALUE 4.00e-32
US-09-864-761-48290
```

```
Query Match 5.0%; Score 311; DB 9; Length 67;
Best Local Similarity 89.6%; Pred. No. 5.4e-14;
Matches 60; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 254 VTKRDEDSLMQLKEEFREYALRREHDAQIVHIAEAGLRISPEQWSSLLYCDLAHKS 313
DB 1 VTKRDEDSLMQLKEEFREYALRREHDSQIVQIAEAGLRAPDQWSSLLYGDOSHSK 60
QY 314 MOSIIDK 320
DB 61 MOSIIDK 67
```

RESULT 6

```
US-10-840-512-196
; Sequence 196, Application US/10840512
; Publication No. US2005012582A1
; GENERAL INFORMATION:
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: MANNING, GERARD
; APPLICANT: CHARYDCZAK, GLEN
```

```
; APPLICANT: GRIGORIEV, IGOR
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-1455
; CURRENT APPLICATION NUMBER: US/10/840,512
; CURRENT FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: 60/469,014
; PRIOR FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 239
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 196
; LENGTH: 2377
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-840-512-196

Query Match 3.3%; Score 207; DB 18; Length 2377;
Best Local Similarity 19.3%; Pred. No. 0.00022;
Matches 216; Conservative 152; Mismatches 42; Indels 326; Gaps 46;

QY 48 KACPFDDTAINT-----DIDLVPVNFALLQLVGAQVDPHQSIKLSNLGEN 92
DB 375 KRASPAKSVIGTPEFMAPEMEYEEKYDESVDYVAFQCMLEMATSEYP----- 421
QY 93 KHYEVAKKCVEDLALYLKPLSGKGVAASLNQSAL-----SRPMQRKULVLVN 139
DB 422 -----YSECQNAAQIYRRVTSQVK-PASFDKVAIPVKEIIEGCIRQNKDERYSIKDLIN 475
QY 140 QOLVEEGRVRAMRAARSLGERTVTTELILQHNQPOOLSANLWAAVRARGCQFLGPAMQEE 199
DB 476 HAFQOETGVRVELAEDDGEKIAIKWLRIEDIKKLK-----GKYKONE 520
QY 200 ALKLVLLALEDGSSALSRKVLVLFVVRLEPRFPQASKTSIGHVVO-----LLYRAS 250
DB 521 AIEP-----SFLERDVPEDVAQEMVESGVCEGDHKTWAKAKDRVS 563
QY 251 CFKVTKRDEDSLMQLKEEFREYALRREHDAQIVHIAEAGLRISPEQWSSLLYCDLAH 310
DB 564 LIK-RKREORQLVREEQEKQKEESSFKQNEQASVS-QAGI-----QLSAASTGIPTA 617
QY 311 KSHMQSIIDKLOSPEFSAKSVQELTIVLQRTCDPANLRLPHLELLANIDPNDAVSPT 370
DB 618 PATSASVSTQVEPEPEADQHQ-----LOYQQPFSISVLVS-DGTIDSGGQS 662
QY 371 WEQLENAMVAVKTVVHGLVDFTONZYRSRGHET-----PQOPNSKYKTSMCRD 418
DB 663 SVFTESRVSSQQTVSYG-----SQHEQAHSTGTAPGHTVSSIOAQSQPHGYVPSSMAQ 716
QY 419 LRQOQGCPRGNTCTFAHSQEELEKYRLNKKINATVTFPLLN--KVGNNNTVTTAGNV 476
DB 717 GQNQOQ-QPSSSLAGVLSQ-----PIHQPOQOQGIQPT----- 747
QY 477 ISVIGSTETTKIVPSTNGISNAENSVSOLISRSTDSTLRALETYKVKVKGVCANGQNAAG 536
DB 748 -----VFSQOAV---QVSLPQAASS-----EG 767
QY 537 PSADSVTENKIGSPKTPVSNVAATSAGPSN--VGTELNSVPQKSSPFLTRVPVPPHSE 594
DB 768 TTAQVSPQSPVSAGTQLPVSQTVAIVQGEPHLPVSTQSPVVPVHSGAHLPMGQPIPTSL 827
QY 595 NIQYQDPDRTOIPIFEVPQY--PQTGYYPPTVPAGVAPCVPRFVRNNSNPVSSLLPPASM 652
DB 828 LPQY---PVSQIPISTPHVSTAQTGFSSVPITMAAGINQ--PLLTASSATASSIPGGSP 882
QY 653 PYADHYST-FSPRDRMNSSPY-----QPPPPQPYGVP-----PVPSPG--MYAPVVD 696
DB 883 VVPNQLPTLLQVNLQSQVHPQLLOPTTVOSIG-IPANLQOAEAGLPSPGGVLYQG-PP 940
QY 697 SRIWRPPMYQRDDIIRNSLPPMDVMHSSVY-QTSLRERYNSLDGYYSVACQPPSEPT 755
DB 941 SR---LPPQYPCDSNIAPSSNVASVCIHSTVLAPPSPMTEALATQGYPTTVQPVESPT 997
QY 756 TVPLPREPCGHLKTSCEBOIRRKRPQMAQYHTQKAPLVSS-----TL 797
```

Db 998 LVPM-GSVGGQVVS-----QPAVSLTQOPPTTSQQAVLESTQVSOAAPPEOT 1046  
QY 798 PVATQPTPSPPLFS-----VDFRADFS-----SVSGTKPEEDHLSHYSPWSCGTIG 845  
Db 1047 PITQSQPTQVPLVTSADSNHSDVAGMSDGNENAFSSGRHGRRTTKHYRK---SVRS 1103  
QY 846 SCINAIDSPKOVIAN-----SNAVLMDLSDGDKRRVHLFETQRTKEDPIIPFSDGP 900  
Db 1104 RSRHETSRLPKILNVKNGDRVBCQLETHNRKMVTFKFDLDGDNPEIATIMVNDNF 1163  
QY 901 IISKWGAISRS-----RTGYHTTDPVQ-----ATASQ 928  
Db 1164 IL-----AIERESFVAQREIIEKADEMLSEDVSEVPEGQGLSELOGKDDYGPFGSKULE 1219  
QY 929 GSATKISVSDYPPYNAVDSRWSYGNESATSAHYVDRDRFIVTDLGHRKHSS---TG 985  
Db 1220 GEFKQPIAVSSMQQIGVPTS-----SLTQVVSAGR-RFIVSPVPSRLRESKVFTS 1271  
QY 986 DLLS-LELOQAKNSLLLOREANALAMQKNSLDEGR 1022  
Db 1272 DISDPVAVASTQAPGNLSHSASSLSLQQAFAFSELKHGQ 1309

## RESULT 7

US-10-092-900A-128  
; Sequence 128, Application US/10092900A  
; Publication No. US20040043382A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Taupier Jr., Raymond J.  
; APPLICANT: Pena, Carol E.A.  
; APPLICANT: Li, Li  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Gusev, Vladimir Y.  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Gorman, Linda  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Gangolli, Esha A.  
; APPLICANT: Vernet, Corine A.M.  
; APPLICANT: Guo, Xiaojia Sasha  
; APPLICANT: Tchernev, Velizar T.  
; APPLICANT: Fernandes, Elma R.  
; APPLICANT: Casman, Stacie J.  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Liu, Yi  
; APPLICANT: Anderson, David W.  
; APPLICANT: Spaderna, Steven K.  
; APPLICANT: Catterton, Elina  
; APPLICANT: Leite, Mario W.  
; APPLICANT: Zhong, Haihong  
; APPLICANT: Alsobrook, John P.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-290C  
; CURRENT APPLICATION NUMBER: US/10/092,900A  
; PRIOR FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: USSN 60/274,322  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: USSN 60/283,675  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: USSN 60/338,092  
; PRIOR FILING DATE: 2001-12-03  
; PRIOR APPLICATION NUMBER: USSN 60/274,281  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: USSN 60/274,191  
; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: USSN 60/325,681  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: USSN 60/304,354  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: USSN 60/279,995  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: USSN 60/294,899  
; PRIOR FILING DATE: 2001-05-31  
; PRIOR APPLICATION NUMBER: USSN 60/287,424  
; PRIOR FILING DATE: 2001-04-30  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 768  
; SEQ ID NO 128  
; LENGTH: 1267  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-900A-128  
  
Query Match 3.3%; Score 203; DB 15; Length 1267;  
Best Local Similarity 19.8%; Pred. No. 0.00016;  
Matches 190; Conservative 116; Mismatches 337; Indels 318; Gaps 46;  
  
QY 423 GCGPRG---TNCCTFAHSQEELEK-----YRLRNKKIN----- 451  
Db 21 GRCPRNSLYSDCIIIEKTVLQKONEGFGFVLRGAKADTPIEEFTPTFAPPALQKLESV 80  
QY 452 -----ATVRTFLLNKVGNNTVTTTGNVISVI--GSTETTGKIVPSTNGI----- 496  
Db 81 DEGGVANOAGLRTGDFLIEVNNENVVKVGHQVVMIRQGNHVLKVVTVTNLDPDDT 140  
QY 497 -----SNAENSVSQILSRSTDTLRALETVKVKGVKANGQNAAGSADSVTENKI 547  
Db 141 ARKKAPPPKRAPTTALTLSKSMTSELELDKPEI-----VPASKPS--RAAENMA 191  
QY 548 GSPPKTPVSNVAATSAGPSN-----VCTELNSVPQKSSPLTRVVPVPPHSENLOYQDPR 603  
Db 192 VEP-----RVATIKORSSRCFPAGSDMNVSGRTLGP-RGRGPTVPPLKSLGQSUYE-R 243  
QY 604 TOIPFEVPQYPTGYVPPPP--TVPAG-----VAPCVPRFVRSNNVPE 644  
Db 244 QGIAVMTPTVGS---PKAPFLIGIPGTWRQKSGITTEERQFLAPPMLKFTFSLMMPD 300  
QY 645 SS--LPPASMPYADHYSTFSPRDMNSSPYQPPPPQ--YG-PVPVPS--GMVAPVYDS 697  
Db 301 TSEDIPPPP-----QSVPPSPPPSPPTTNCPSKPTPRVYGTIKPAFNQ 344  
QY 698 RRIWRPPMYQRDDIIRSNSLPPMDVMHSSVYQTSLEERYNSLDG--YYSVACOPPSEPRT 755  
Db 345 NSAAKVSPATRSDDTVATM-----MREKGMVFRRELDRY-SLDSSELYSRNAGPQANFRN 397  
QY 756 -TVPLPREPCGHLKTSCEQI-----RRKPDQMAQYHTQKAPLVSSSTLPVAT----- 801  
Db 398 KRQMPENPYSEVGKIASKAVVYPAKPARKGMVLVKQSNVEDSPKTCSTPIPTIIVKEP 457  
QY 802 -----QSPTPPSPLFSVDPRADPSESVSQTKFEDHLSHYSPWSCGT 843  
Db 458 STSSSGKSSQSGSMEIDPOAPEPPSQL-----RPDESLSVTS-----SPFAAAI 500  
QY 844 IGSCINAIDSEPKDVIANSNA-VLMDLDSG-----VKERVHLPETQRTKED-- 891  
Db 501 AG-----AVRDREKREARENSPAFLSTDLGDEVDVGLGPPAPRTKPSMFPSEEGFADESDA 556  
QY 892 -----PIIPFSDGPIISKWGAISR-----SSRTGYHTTDPVQATASQSA-- 931  
Db 557 EQLSSPMPSATPREPNHFVGAEASAPGEAGPLNSTSKAQGPSSPAVPASSGTAGP 616  
QY 932 -----TKPISV-----SDYVP-----YVNV-----AVD 948  
Db 617 GNVVHPLTGRLLDPPSSPLALALSARDAMKESQSGPKGEAPKADLNKPLYIDTKMRPSLD 676  
QY 949 SRWSSYGNENAT--SSAHYVERDRPIVTDLSGRKHSHSTGDL--SLELOQAKNSLLLORE 1005  
Db 677 AGFPTVTRQNRGRLRRQETENKYETDLGRDRKDKGNMLDIDMTSQQKSAGLLMVT 736

```

Qy 1006 ANALAMQKWNLSDEGRHLTLNLLSKETIERNGELQSDYTEDATDTK-----1052
      :      ||      :      :      :      :      :      :
Db 737 VDATKLDNALQEED-----KAEVEMKPDSSPEVPEGVSTEGALQISAAPEPT 787
      :      :      :      :      :      :      :      :
Qy 1053 -PDRDI-----ELE-----LSALDTEPDGQSPETIHEILDIGLIGSSQNDQL 1094
      :      :      :      :      :      :      :      :
Db 788 VPGRTIVAVGSMEEAVILPFRIPPPPLASVDLDEDFITTEPLPLSFANSFIDPDRAA 847
      :      :      :      :      :      :      :      :
Qy 1095 NGMAVENGHPVQHQKEPPKQKQSLGEDHVHLEEQKTLIPVTSFCSPQLPVSISNASCL 1154
      :      :      :      :      :      :      :      :
Db 848 SVPAL-----SDLVKQKSDTQSPSLNSSQ-----PTNSADSKK-PASLSN--CL 890
      :      :      :      :      :      :      :      :
Qy 1155 P 1155
      :      :
Db 891 P 891

RESULT 8
US-10-732-923-6679
; Sequence 6679, Application US/10732923
; Publication No. US20050108791A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52796)C
; CURRENT APPLICATION NUMBER: US/10/732,923
; CURRENT FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: 10/310,154
; PRIOR FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 24149
; SEQ ID NO 6679
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Candida tropicalis
US-10-732-923-6679

```

```

RESULT 9
US-10-334-143-77
; Sequence 77, Application US/10334143
; Publication No. US20040009549A1
; GENERAL INFORMATION:
; APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL
; TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD
; FILE REFERENCE: 038602/1543
; CURRENT APPLICATION NUMBER: US/10/334,143
; CURRENT FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: 60/343,169
; PRIOR FILING DATE: 2001-12-31
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 77
; LENGTH: 3586
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-334-143-77

```

Db 1689 SLLQPKGPPEPAKQGGADTLSTQTMGKSGHNLVEERTVNLKPAKGTGVKVIDRQ 1748  
QY 363 NPDAVSPTWEOLENA-----MVAVKTVVHGLVDFIQY-SRKGHETPQOPNSK 410  
Db 1749 NDALSKSLRRLSNHHKSNVLESGDKTGTWDTTGEQHLRDEYMSRQLTSTSVKNNLT 1808  
QY 411 YKTS--MCRDLR-----QQGCGPRGNCTF--AHSQEELEKYLRLNKKINATVTRTP 458  
Db 1809 TKESDRAVELKDDVFNISQAGTKVGKQOTVELRNDHQWEGFHKSPKTKNLIK--- 1865  
QY 459 LNKVGNNVTYTTAGNVLSVIGSTETTKIYVSTNGISNAENSVQSLSRSTDSTLRAL 518  
Db 1866 -----ILTDTOSS-KPSPHQVPMVPGVGTVDLSGDFQKQTLKQ 1904  
QY 519 ETV---KKVGKVGANGN-----AAGPSADSVTENKIG----- 548  
Db 1905 ETYSNKDKKKNNINLOPMWQLLPVEQDTSNVTMKVSEKSHNTFKATNKKRETVDHLKS 1964  
QY 549 -----SPPKTPVSNVAATSAGSNVGTSLNSVPQKSSPFLTRVVPVPHSENIQVQDP 602  
Db 1965 QDFLTKTNTSTGLKMAERSLNPINFNPE-NNVKESECLPPLPPSPPPPPS-----NA 2016  
QY 603 RTQIPPEVQYQYPTGYPPPTPVAGVAPCVPRFVS-----NNVPSSLPASM 652  
Db 2017 SSIEFFLP-----PPPLM---MFPEKNGFLPSLSTEKIAEFESFPGLPLPP--- 2062  
QY 653 PYADHYSTSPDRMNSPVQPPQPYGVPVPSGMVAPVYDSRRIRWPPMPYQDDII 712  
Db 2063 PVDEKS-----BRESSMFLPPP-----PPTPS-----QXPAHL 2093  
QY 713 RNSLPPMDVMHSS--VYQTSLEERYNSLDGYYSV-----ACQPSSEPTTVPLPREP 764  
Db 2094 LSSAPE---KHSGDFMQYVSOKEASNSQNSQAKIITKGTGLVLP-----TLPKPLP- 2144  
QY 765 GHUK-----TSCEBQIRKPDQ-----WAQVHTO-KAPLVSTSLPVAT 801  
Db 2145 KHKDKNDNDFSPKVELATSLSDMECKITTSKQKVMVMTSSEHTETKQNVISKSLDERK 2204  
QY 802 QSTPPSPPLSFVDFRADFSVSGTKFEEDH-----LSHSPWSGGTIGSCINAIIDSP 855  
Db 2205 Q-----LSIDANCLSHVPTGAPRKQIAPLIKSHSPFESSG-----QQNP 2247  
QY 856 KOVIANSNAVIMDLSDGVKRRVHLFTQRTKTEEDPIIPFSDGPIISKWGAISRRTG 915  
Db 2248 KPYMRKEFTPLM---TAEKYRQKEIEKQKE-----SSYNIKVTQSQN 2291  
QY 916 YHTD-----PVQATASQGSATKPISVSDVVPVYVNAVDSRWSSYGN-----AT 959  
Db 2292 QHTEVEKEMLQKT-----NEEVSLSGIDSECTVVQPSQSNARILGVCSNDQLST 2345  
QY 960 SSAHYVERDRFIVTDLSGHKKHSTGDL--SLELOQAKNSLLO--REANALAMQOK- 1014  
Db 2346 TSPETVAAGKLHVLAASEDKMKXKEVLQSSHDIMQSKACEIKOSHOBECSTQOQOKK 2405  
QY 1015 ---WNSLDEGRHLTLNLLSKEIELRNGELQSDYEDATD-----TKPDRDIEEL--S 1062  
Db 2406 YLEQLHLQPSKIPSPNFVKVTKILPT---LDHTLNETHSYSHKQOSEIDVQTKKQ 2461  
QY 1063 ALDTDEPDGQSEPIEELDIOLGISSQNDQLNGMAVENGHVPQOQKEPPKOKK----- 1117  
Db 2462 YLTKTKTEASTE-----CSHKQSL-----AERYQLPKKEKRVTVQL 2498  
QY 1118 --OSLGEDHVLBEQKTLVPTSCFQPLPVSIASNACSLPITTSVSNAGNLIL---KTHVM 1172  
Db 2499 PTESIQKNQ---EDKLMVP-----RKQREFSGDRGLPGSEBKQGPSPMIGRKEERLI 2550  
QY 1173 SEDKNDPLKPVANGKVV 1189  
Db 2551 TERKHEHLKNSAPKVV 2567

RESULT 10

US-10-437-963-126940

; Sequence 126940, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 126940  
; LENGTH: 1248  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(1248)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_29438C.1.pap  
; US-10-437-963-126940

Query Match 3.2%; Score 196.5; DB 16; Length 1248;

Best Local Similarity 21.6%; Pred. No. 0.00044;

Matches 148; Conservative 84; Mismatches 235; Indels 219; Gaps 35;

QY 436 SOEELEKYLRLNKKINATVTRTPPLNKGNNVTVTITAGNVISVIGSTETTKGIVPSTNG 495  
Db 408 SOGNDXLLMLGPRSS-----ILKK---NLTLVLSSPVQQLIDSAVQIAQSA 458  
QY 496 ISNAENSVQSLSRSTDSTLRALFTVKYKGVKANGQNAAGPSADSV---TENKISGPK 552  
Db 459 SKSAENSGSQTVPNGEPSPL---TV-----NNSASTASIALCTPPPLPPPP 503  
QY 553 T-----PVSNAATSAGSNVGTSLNSVPQKSSPFLTRVVPVPHSENIQVQDPRTQIPF 608  
Db 504 TVSLAPVSPILINTSTSIISVLSIMPSPSTAASPL-PP-----PL 547  
QY 609 EVPPQYPTGYVPPPTPVAGVAPCVPRFVRNNVPE-SSLPPASMPYADHYSTSPDRDM 667  
Db 548 PPPLKPESTVMFPFPPPPPPPPPPPPPP--VSSFSPQPPPPPPSASP-GL 593  
QY 668 NSSPYQPPPPQYGVPPVPVPSGMYAFVYDSRRIRWPPMYQRDDIIRSNLPPMDVMHSSV 727  
Db 594 QASPVPPPPPP--PPPMIP-GMKTTP-----TTP-----PPPP----- 624  
QY 728 YQTSLEERYNSLDGYYSVACQPPSEP-----RTTVLPREP-----CGHLKTSCEEQI 775  
Db 625 -----AAGQOAPAVPPPPPPPPPPPPPPVGMQTRIPPPPPSPQNSLVSFPSTS 674  
QY 776 RKPKDQWQYHTQKAPLVSTLP-----VATQSPTPSPPLSVDFRADFS----- 822  
Db 675 KRIPPP-PPPPQTSLSVSS-LPSSRKGNDAAPRPPPPPPPLYSRSHVTSASAPAPP 732  
QY 823 ---VSGTKFEEDHLSHYSP-----WSCGTIGSCINAIIDSEP-----KD 857  
Db 733 LPPPKLVGASKSQEQMITWPPPPPPPKNSNSLPSKGNVSVSSPPPPPTFSFGAKD 792  
QY 858 -VIANSNAVIMDLSDGVKRRVHLFTQRTKTEEDPIIPFSDGPIISKWGAISRRTGY 916  
Db 793 RSTARSRS-----PRSLRPNQSKRT-----PLKPL-----HMVKVSR----- 825  
QY 917 HTTDPVQATASOGSATKPISVSDVVPVYVNAVQ-SRWSSYGNBATSAAHYVERDRFIVTDL 975  
Db 826 -----ATOGSLWAEQKSDASRTPEIDISELESLSFVAMPN---MEEXR----- 867





```
Query Match 3.1%; Score 195; DB 15; Length 2135;
Best Local Similarity 19.1%; Pred. No. 0.0013;
Matches 254; Conservative 169; Mismatches 528; Indels 378; Gaps 56;

QY 114 GKGVASLNQASLSRPMQKRLVTLVNCQLVEBGRVRAARSLGERTVTELLIHOQNP 173
DB 720 GQFSSSLTGVSSQPIHQPOQGIQOQTAPPOQVQISQTSSEATTAQVSPQAP 779
QY 174 QQLSANLWAAVRARGCOFLGPA-----MOEALKVLLALEDGSALSRRKVLVLFVORLE 228
DB 780 QVLQVQASQSTQGVQVAPAEVPAVQATQPTTLASSVDSAHSD---VASGMSDGN 836
QY 229 PRFPOAKTSIGHVOLLVYRASKVKTKRDESSLMQK-----BEFRSYEARREHDA 282
DB 837 ENVPSSSGRHEGRTTKRHYKRS-VRSRSHKTSRPLKAILNVNKGDRVVECOLETHNR 895
QY 283 QIVHIAMEAGLRISQWSSLLY-GDLAKSHMQSIIDKQSPESPAKSVQELITVLQRT 341
DB 896 KMTVTKFDLD-GDNPEIATIMVNDIFILAIBRESFVDQVREIIKADSEMLSEDSVPE 954
QY 342 GPANLNLRLPHLELLANID-----PNPDVSPPTWEOLENAMVAVKTVV 385
DB 955 GQOG-----LESQKDDYGFSGSKLEGEFKQIPASSMP--QQIGIPTSSLTQVV 1004
QY 386 HGLVDFIONYSRKGHETPOQPN-----SKYKTMCDRLRQGGCPRGNTCTFAHSQ 437
DB 1005 H-----SAGRFVSPVPSRLRESKVPSEITDTVAASTAQSPG-----MNLSHA 1051
QY 438 BELEKYR-----LRNKKI-----NATVRTFP-----LNKVGNNVTVTTA----- 473
DB 1052 SLSLQQAQSELRAQMTGPNATPNFSGHTGTPVPVPPFLSSIAAGVPTAAATAPVA 1111
QY 474 -----GNVLSVSGTETTKIVPSTNGISNAENS-----VSQILSR 509
DB 1112 TSSPPNDISTVIOSEVT---VPTREGIAGVATSGVTTSGGLPIPPVSESPLVSSWSS 1168
QY 510 STDSTLRALETVKVKGVKAN-----GQNAAGPSAD-SVTENKIG-----SPKPTVSNV 558
DB 1169 ITIPAVVSISTSPSLOVTSISEIVVSTALVPSVTVSATASAGGSTATGPKPPAVV 1228
QY 559 AATSAGPSNVGTLSNV-----PQSSPFLTRVVPVPHSENIQYFQ 600
DB 1229 SQQAAGSTTVGATLTSVSTTSPSTASQSLQSSSTSTPTLAETVVVSAHSLD-KTSH 1287
QY 601 DRTQITPFPVQPTGYVPPPP-----TVPAGVPCV-PRFRSNVNPVPSLPPASMPY 654
DB 1288 SSTTGLAFSL-SAPSSSSFGAGVSSYISQPGGLHVLVPSVIASTPILPQAAGTSTPL 1346
QY 655 ADHYSTFSRDRMNSSPYPPPPQPYGPPVPPVPSGM-----VAPYDGR 698
DB 1347 -----LPQVPSIPPLVQPVANVPAVQOOLIHQSQPQALLPNQPHTHCEVDS 1394
QY 699 RIWRPPMYQRDDI-----IRS--NSLPPMDVMHSSV-YQTSILRERYNSLDGYIVA--- 746
DB 1395 T--QPKAPGDDIKTLEKLSLFSHSSSGAQSLSVSTSTVPGITAVAP 1452
QY 747 -----CQPPSE-PRTTVPLPREPC---GHLKTSCEBQIRR-KPDQWAOYHTQKAP 791
DB 1453 SKLLTSTTCLPPTNLGLTVALPVTVPVTPQGVSTPVSTTSGVKPGT----APSKPP 1508
QY 792 LVSS--TLPVATQSTPPSPPLFSDVPADPSESVSGTKFEEDHLSHVSPWSCCTGSCINA 850
DB 1509 LTKAPVLPVGTLEP-----AGTLFS--EQ 1530
QY 851 IDSEPKDVIANGNAVLMDLDS-----GDVKRRVHLFETQRTK----- 888
DB 1531 LPPFPGLSTQSOQLEDLDAQLRRLTSPMITVTSVAGPVSNMAAPTATEAGTQPKGV 1590
QY 889 -----BEDPIIFSDGPIISKWG-----AISRSTGYHTDPVQATASQGSATKPISVSD 939
DB 1591 SQVKGPVLATSSGAGVFKMGRFQVSAADGAQKKGKNSDEAKSVHFESSSTSESSVLSS 1650
```

```
QY 940 YVPYNAV-----DSRWSSYGNBATS-SAHYVERDRFIVTDL-----GH 978
DB 1651 SSPESTLVKPEPNGITIPGISSDVPESAHTTASEAKSDTQPTKVGRFQVTTANKVGR 1710
QY 979 RGHSGTGLLS-----LELQAKSNSLLQREANALAMQOKWN----- 1016
DB 1711 FSVSTKEDKTDTKKEGPVSPFMDLEQAVLPVPIPKKEKPELSEPSHLNGPSSDPEAA 1770
QY 1017 -----SLDEGR---HUTLMLLSKEIELRN-----GELOSDY--TEDATDTKPDRIEELS 1062
DB 1771 FLSDRDVDDSGSPHSPHQLSSKSLPSQNSLSNSFNSSYMSNDESIEDDLKLELR 1830
QY 1063 ALDTEPDQSPSEPIEILDIQIGISSQNDQLLN-----GMAVENGHVPQHQKPEPPKQ 1115
DB 1831 RL-----RDKHLKEIQDLSRQKHEIESLYTKLGKVPVAVIIPPAAPLSGRRRPTKS 1883
QY 1116 K-----KQSLGEDH-----VILEQKTLPTVTSFQSPPLVPSISNASCLPI 1156
DB 1884 KGSKSSRSSSLGNKSPQLSGNLSGQSAASVLHPQOQLHPPGN-----IPESQOQLQLPL 1938
QY 1157 TTSVSAGNL 1165
DB 1939 KPSPSSDNL 1947

RESULT 13
US-10-362-892-9
; Sequence 9, Application US/10362892
; Publication No. US20040038881A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; BANDMAN, Olga
; APPLICANT: NGUYEN, Dannie B.; WALIA, Narinder K.
; APPLICANT: HAPALIA, April J.A.; YAO, Monique G.
; APPLICANT: GANDHI, Ameena R.; GURURAJAN, Rajagopal
; APPLICANT: DING, Ji; PATTERSON, Chandra S.
; APPLICANT: YUE, Henry; BAUGHN, Maria R.
; APPLICANT: TRIBOULEY, Catherine M.; THORNTON, Michael B.
; APPLICANT: ELLIOTT, Vicki S.; LU, Yan
; APPLICANT: ISON, Craig H.; AU-YOUNG, Janice K.
; APPLICANT: TANG, Y. Tom; AZIMZAI, Yalda
; APPLICANT: BURKILL, John D.; MARCUS, Gregory A.
; APPLICANT: ZINGLER, Kurt A.; LU, Dyung Aina M.
; APPLICANT: LAL, Preeti G.; RAMKUMAR, Jayalaxmi
; APPLICANT: WARREN, Bridget A.; KEARNEY, Liam
; APPLICANT: POLICKY, Jennifer L.; THANGAVELU, Kavitha
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PF-0209 USN
; CURRENT APPLICATION NUMBER: US/10/362,892
; PRIOR FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,873
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/240,542
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 2135
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

|  |      |  |      |
|--|------|--|------|
| Db   | 1591 | SQVKEGVLATSSGAGVFKMGRQVSVAAQCAQKGNKSEDAKSVHFSESTSESSVLSS     | 1659 |
| Qy   | 940  | YVPYNAV-----DSRWSSYNGNEATS-SAHYVERDRFTVTDLS---GH             | 978  |
| Db   | 1651 | SSPESTLVKPEPNGIITPGISSDPESAHTKTASEAKSDTGQTKVGRFOVTTANKVGR    | 1710 |
| Qy   | 979  | RKHSSTGDLSS-----LELOQAQNSLLLOREANALAMQOKWN-----              | 1016 |
| Db   | 1711 | FSVSKTEDKITDTKKGEPVASPPFMDLEQAVLPAVIPKKEPELSFSLHNGPSSDPEAA   | 1770 |
| Qy   | 1017 | -----SLDEGR--HLTILNLSKEIERN-----GELQSDY-TEDATDTPKPRDIELELS   | 1062 |
| Db   | 1771 | FLSRDVEDGSGSPHSQJSSKSLPQNSLSQSLNSFNSSYMSDSNEDIEDDLKLELR      | 1830 |
| Qy   | 1063 | ALDTDEPDGQSEPIEBILDTQLGISSQNDQLLN-----GMAVENGHVPQOQHKEPPKQ   | 1115 |
| Db   | 1831 | RL-----RDXHLKEIQDLSQKHEIESLYTKLGKVPVAVIPPAAPLGRRRRTKTS       | 1883 |
| Qy   | 1116 | K-----KQSLGSDH-----VILEBQTKILPVTSCFSQPLPVSISNASCLPI          | 1156 |
| Db   | 1884 | KGSKSRSSSLGNKNSPQLSGNLSQSAASVLHPQOTLHPPGN-----IPESGQNQLLOPL  | 1938 |
| Qy   | 1157 | TTSVSAGNL 1165   |      |
| Db   | 1939 | KPSPSSDNL 1947   |      |
| RESULT 14  |      |  |      |
| US-10-196-935A-2   |      |  |      |
| ; Sequence 2, Application US/10196935A                                       |      |  |      |
| ; Publication No. US20030082720A1  |      |  |      |
| ; GENERAL INFORMATION:   |      |  |      |
| ; APPLICANT: Lifton, Richard P   |      |  |      |
| ; APPLICANT: Wilson, Frederick H   |      |  |      |
| ; APPLICANT: Choate, Keith   |      |  |      |
| ; APPLICANT: Ishikawa, Kazuhiko  |      |  |      |
| ; APPLICANT: Nelson-Williams, Carole   |      |  |      |
| ; TITLE OF INVENTION: COMPOSITIONS METHODS AND KITS RELATING TO TREATING AND |      |  |      |
| ; TITLE OF INVENTION: HYPERTENSION   |      |  |      |
| ; FILE REFERENCE: 044574-5113  |      |  |      |
| ; CURRENT APPLICATION NUMBER: US/10/196,935A                                 |      |  |      |
| ; CURRENT FILING DATE: 2002-10-25  |      |  |      |
| ; PRIOR APPLICATION NUMBER: US 60/306,084                                    |      |  |      |
| ; PRIOR FILING DATE: 2001-07-17  |      |  |      |
| ; NUMBER OF SEQ ID NOS: 6  |      |  |      |
| ; SOFTWARE: PatentIn version 3.1   |      |  |      |
| ; SEQ ID NO 2  |      |  |      |
| ; LENGTH: 2382   |      |  |      |
| ; TYPE: PRT  |      |  |      |
| ; ORGANISM: Homo sapiens   |      |  |      |
| US-10-196-935A-2   |      |  |      |
| Query Match 3.1%; Score 195; DB 14; Length 2382;                             |      |  |      |
| Best Local Similarity 18.5%; Pred. No. 0.0015;                               |      |  |      |
| Matches 235; Conservative 167; Mismatches 506; Indels 362; Gaps 49;          |      |  |      |
| Qy   | 48   | KACPFDOATAINT-----DIDVLPNFALLQLVGAQVDPHQSIKLSNLGEN           | 92   |
| Db   | 375  | KRASFAKSVGTGTFEFWAPENYEEKYDSVDVYAFGCMLEMATSEYP-----          | 421  |
| Qy   | 93   | KHYEVAKKCVEDLALYLKPLSGKGVASLNQSAL-----SRPMQRKLVTLVN          | 139  |
| Db   | 422  | -----YSECQNAQIYPRVTSVGVK-PASFDKVAIPEVKEIIEGICRQNKDERYSIKDLN  | 475  |
| Qy   | 140  | COLVVEEGRVRAARAARSLGERTVTTELILQHNQPLQSANLWAARVARGCOFLGPAMQEE | 199  |
| Db   | 476  | HAFQOEBTGVGRVELAEDDDGEKIAIKLWLRIEDIKKLK-----GKYKDNE          | 520  |
| Qy   | 200  | AKKLVLALLEDGSAISRKVLVLFVQFLEPRFPQASKTSIGHVVQLLYRASCFCFKVKRDE | 259  |
| Db   | 521  | AIFP-----SFDLERDVPEDVAQEMVESGYVCEGDHKTWAKAIKOR--             | 561  |



|    |      |                 |                     |                    |                        |      |
|----|------|-----------------|---------------------|--------------------|------------------------|------|
| Qy | 200  | ALKLVLLALEDGSA  | SRKVLVLFVVQRLPRFPQ  | ASKTSIGHVOLLVYRASC | FKVTKRDE               | 259  |
| Db | 521  | ALIEF-----      | -----SFDIERVDPED    | VAQEWESGYVCEGDH    | KMAKADR--              | 561  |
| Qy | 260  | DSLSMLQKEFRSYE  | ALRREHDAQIVHIAE     | AGLRISPEQWSSLYL    | GDLIAHSHMQSIID         | 319  |
| Db | 562  | -VSLIKRKEQR-    | -QLVREEQEK-         | -----KQESSLKQ      | QVEQSSASQTGIK          | 604  |
| Qy | 320  | KLQSPESFAKSVQEL | TIVLQRTGDP-----     | ANLRLRPHLELLAN     | IDPNPDVSPTW            | 371  |
| Db | 605  | QLPSASTGTPASTT  | SASVSTQVEPEEPA      | DQHQLQYQOPSISVL    | --DGTVDSGQGSS          | 662  |
| Qy | 372  | EQLENAMVAVKTVV  | HGLVDFIQNYSRKGHET   | -----POPOENSKYK    | TSMCRDL                | 419  |
| Db | 663  | VFTESRVSSQQT    | VSIG-----SQHEQA     | HSTGTVPCHIPSTV     | QAOSQPHGYVPPSSVAQG     | 716  |
| Qy | 420  | RQGGCPRGTCNCT   | FAHSQBELEKYLRLN     | KKINATVTRTFLPLN    | KVGVNNTVTTTGNVISV      | 479  |
| Db | 717  | QSQGQ-PSSSSLT   | GVSSSQPIQHPO-----   | QOQGIQQTAPPQ-      | -----                  | 752  |
| Qy | 480  | IGSTETTKIVPSTN  | GISNAENSQGISRST     | DSLRALETVKVKVG     | KANGCONAAGPSA          | 539  |
| Db | 753  | -----QTVQYSLQ-  | TSTSSEAT-----       | TAQPVSQPOA         |                        | 779  |
| Qy | 540  | DSVTENKIGSPKTP  | VSUNVAATSAGPSN--    | VGTELNSVPQKSPP     | FLTRV-----PVYPP        | 591  |
| Db | 780  | PQVLP-QVSAGQ    | LPVSPVPTIQGEPQ      | IPVATQSPVVPVH      | GAHFLPVGQPLPTPLP-      | 837  |
| Qy | 592  | HSENIQYFQDPR    | TQIPFEVPOX--POT     | GYPPPTVPAGVPC      | VPFRFVRNNVPSSLP-       | 648  |
| Db | 838  | -----QYVVSQ     | PISTPHYSTAGTGF      | SSLPITWAAGITQ--    | PLLTILASSATTAAPG       | 887  |
| Qy | 649  | -----PASMPY     | ADHYSTFSRDRMNSPY    | -----QPPP-----     | POPYGPVPPVPSGMYAPV     | 694  |
| Db | 888  | VSTVWPSQLP-     | -----TLQPV          | TQLPSQVHPQLLQ      | PAVQSMGIPANLQAAE       | 942  |
| Qy | 695  | YDSRRIRWPPMY    | QORDIIRSNSLPPMD     | VHSSVYQTSLRERY     | NSLDGYYSVACQPPSEPR     | 754  |
| Db | 943  | YOGFPRLFPQ      | PGDSNIAPSSNVA       | SCVTHSVLSPMP       | TEVLATPGYPTTVVQPVESN   | 1002 |
| Qy | 755  | TTVPL-----      | PREPCGHL-----       | KTSCEEQIRRRKPD     | QMAQVHTQKAPLVSSTLP     | 802  |
| Db | 1003 | LLVPMGCVGQ      | VQVQSGSLAQAPT       | TSSQAVLES-----     | TQGVSOVAPAEPPAVA       | 1054 |
| Qy | 803  | SP--TPPSPL-     | FSVDF-----RAD       | FSVSGTKFEEDHL      | SHYSPWSCGTIGSCINAI     | 851  |
| Db | 1055 | QPOATQPTTLAS    | SVDSAHSDVASGMS      | DGNENVPSSGRH       | EGRTTKRHYRKSVRSR       | 1114 |
| Qy | 852  | DSEPKDVIAN-     | -----SNVLM          | DLDSGDVKRRVHLE     | TQRTKEEDPIIPFSDG       | 906  |
| Db | 1115 | TSRPLRLNVN      | KNKGRVVEQCLETH      | NRRKMTFKFDLDG      | NPEEITATIMVNNDFIL----  | 1170 |
| Qy | 907  | AISRSS-----     | RTGYHTTDPVQA-----   | TASQ---GSATKP      |                        | 934  |
| Db | 1171 | AIERESFVQVREI   | IEKADEMILSEDS       | VEPEGQGLSLOGK      | DDYCFSGSKLEGEFKQ       | 1230 |
| Qy | 935  | ISVSDYVPYVNA    | VDSRWSSYGNEAT       | SSAHYVERDRFIV      | TDLSGHRKHSS---TGDLL-SL | 990  |
| Db | 1231 | IPASSMPQIGI     | PTS-----SLTQ        | VHSAGR-RFIVSP      | VPESRLRESKVPFSEITDV    | 1282 |
| Qy | 991  | ELQOAKSNSLL     | LOREANALAMQOK       | WNSLDEGRHLTLN      | LLSKLELNGELQSDY        | 1050 |
| Db | 1283 | AASTAQSPGNL     | SHSASSLSLOQAFS----- | ELFRAQM----        | TEGPNT                 | 1321 |
| Qy | 1051 | TKPDRD-----     | -----IELE           | SALDTPDQGEPIE      | ELD--IQLGISQNDQ        | 1092 |
| Db | 1322 | APNFSHTGTF      | PPVPPFLSSTAGV       | ETTAATAPVAT        | SPNDISTSVIQSEVTVPT     | 1381 |
| Qy | 1093 | LLNGMAVENGH     | FPVQHQKPPKQK        | QSLGEDHVILEE       | KTILFVTSFCFQPLFV       | 1152 |
| Db | 1382 | GIAGVATSTG      | VTSGGLPIPP-----     | VSESPVLSVSV        | SITIPAVVSI             | 1428 |

Search completed: August 4, 2005, 01:53:32  
Job time : 181 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: August 4, 2005, 01:25:48 ; Search time 179 Seconds  
(without alignments)  
2573.361 Million cell updates/sec

Title: US-10-619-992-2

Perfect score: 6215

Sequence: 1 MPVQAQWTEFLSCPICYNE.....MSEKNDFLKPVANGKMWNS 1191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Genesecp1980s:\*
- 2: Genesecp1990s:\*
- 3: Genesecp2000s:\*
- 4: Genesecp2001s:\*
- 5: Genesecp2002s:\*
- 6: Genesecp2003as:\*
- 7: Genesecp2003bs:\*
- 8: Genesecp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match % | Length | DB ID | Description         |
|------------|--------|---------------|--------|-------|---------------------|
| 1          | 6215   | 100.0         | 1191   | 5     | AAE18608 Human cel  |
| 2          | 6207   | 99.9          | 1191   | 5     | AAE18609 Human cel  |
| 3          | 6206   | 99.9          | 1191   | 5     | AAE18613 Human cel  |
| 4          | 6206   | 99.9          | 1191   | 5     | AAE18612 Human cel  |
| 5          | 6205   | 99.8          | 1191   | 5     | AAE18611 Human cel  |
| 6          | 6205   | 99.8          | 1191   | 5     | AAE18610 Human cel  |
| 7          | 5483   | 88.2          | 1048   | 7     | ADb80934 RING-SH c  |
| 8          | 4316   | 71.1          | 869    | 4     | AAm40354 Human pol  |
| 9          | 4066.5 | 65.4          | 812    | 4     | AAm40355 Human pol  |
| 10         | 2665.5 | 42.9          | 1125   | 7     | ADc37580 Human nov  |
| 11         | 2577.5 | 41.5          | 1109   | 7     | ADf55448 Human nov  |
| 12         | 2106   | 33.9          | 421    | 4     | AAg73504 Human gen  |
| 13         | 1593   | 25.6          | 306    | 4     | AAm42140 Human pol  |
| 14         | 1593   | 25.6          | 306    | 4     | AAm42141 Human pol  |
| 15         | 1538   | 24.7          | 819    | 4     | ABb66041 Drosophill |
| 16         | 1492.5 | 24.0          | 296    | 4     | AAg73505 Human gen  |
| 17         | 611    | 9.8           | 522    | 7     | ADb65511 Human pro  |
| 18         | 490    | 7.9           | 95     | 5     | ABp64303 Human ORF  |
| 19         | 311    | 5.0           | 67     | 5     | ABg46802 Human pep  |
| 20         | 267    | 4.3           | 276    | 4     | ABg11134 Novel hum  |
| 21         | 206    | 3.3           | 1556   | 8     | ADq97964 Human can  |
| 22         | 204.5  | 3.3           | 127    | 3     | AAb42232 Human ORF  |
| 23         | 203    | 3.3           | 1267   | 5     | ABu65104 Human NOV  |
| 24         | 199.5  | 3.2           | 1994   | 8     | ADr18912 Human muc  |
| 25         | 198.5  | 3.2           | 3586   | 7     | ADk40970 Novel hum  |

|    |       |     |      |   |          |                    |
|----|-------|-----|------|---|----------|--------------------|
| 26 | 198.5 | 3.2 | 3586 | 8 | ADR15689 | Adr15689 Kinase 72 |
| 27 | 197.5 | 3.2 | 867  | 8 | ADS88438 | AdS88438 Human pro |
| 28 | 197   | 3.2 | 713  | 8 | ADR97079 | Adr97079 C. albica |
| 29 | 195.5 | 3.1 | 123  | 4 | AAm21438 | AAm21438 Peptide # |
| 30 | 195.5 | 3.1 | 123  | 4 | ABb43777 | ABb43777 Peptide # |
| 31 | 195.5 | 3.1 | 123  | 4 | AAm37681 | AAm37681 Peptide # |
| 32 | 195.5 | 3.1 | 123  | 4 | ABb26719 | ABb26719 Protein # |
| 33 | 195.5 | 3.1 | 123  | 4 | AAm77509 | AAm77509 Human bon |
| 34 | 195.5 | 3.1 | 123  | 4 | AAm64739 | AAm64739 Human bra |
| 35 | 195.5 | 3.1 | 123  | 4 | ABg59147 | ABg59147 Human liv |
| 36 | 195.5 | 3.1 | 123  | 5 | ABq46532 | ABq46532 Human pep |
| 37 | 195   | 3.1 | 1650 | 8 | ABm81216 | ABm81216 Tumour-as |
| 38 | 195   | 3.1 | 2135 | 5 | AAe21714 | AAe21714 Human PKI |
| 39 | 195   | 3.1 | 2382 | 6 | ABp71619 | ABp71619 Human WNK |
| 40 | 195   | 3.1 | 2382 | 7 | ADJ69598 | ADJ69598 Human hea |
| 41 | 195   | 3.1 | 2388 | 8 | ADs11119 | ADs11119 Human the |
| 42 | 194   | 3.1 | 1267 | 8 | ADN61859 | ADn61859 Human nov |
| 43 | 194   | 3.1 | 2185 | 7 | ABO44398 | ABo44398 Novel hum |
| 44 | 194   | 3.1 | 2245 | 7 | ABO44382 | ABo44382 Novel hum |
| 45 | 194   | 3.1 | 2322 | 7 | ABO44397 | ABo44397 Novel hum |

## ALIGNMENTS

RESULT 1  
AAE18608  
ID AAE18608 standard; protein; 1191 AA.  
XX AC AAE18608;

XX DT 17-MAY-2002 (first entry)

XX DE Human cell surface DNA receptor (DNA-R) #1.

XX KW Human; cell surface DNA receptor; DNA-R; inflammation-associated disease;  
otitis media; septic arthritis; drug screening; gene therapy; virucide;  
antitumour; antibacterial; immunosuppressive; chromosome 9q34;  
auditory.  
XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
FT Domain 14. .50  
FT Misc-difference /note= "C3HC3D ring finger domain"  
FT Misc-difference /note= "N-linked glycosylation"  
FT Misc-difference 394  
FT Domain /note= "N-linked glycosylation"  
FT Misc-difference 416. .435  
FT Misc-difference /note= "Zinc finger domain"  
FT Misc-difference 430  
FT Misc-difference /note= "N-linked glycosylation"  
FT Misc-difference 451  
FT Misc-difference /note= "N-linked glycosylation"  
FT Misc-difference 466  
FT Misc-difference /note= "N-linked glycosylation"  
FT Misc-difference 468  
FT Misc-difference /note= "N-linked glycosylation"  
FT Region 549. .809  
FT Domain /note= "Proline rich region"  
FT Domain 1133. .1171  
FT Misc-difference /note= "Transmembrane domain"  
FT Misc-difference 1150  
FT Misc-difference /note= "N-linked glycosylation"

XX WO200210392-A2.

XX PD 07-FEB-2002.

XX PF 01-AUG-2001; 2001WO-US024351.

XX PR 01-AUG-2000; 2000US-0222624P.

XX (UYOR-) UNIV OREGON HEALTH SCI.  
PA (USGO ) US DEPT VETERANS AFFAIRS.  
PA (TARG-) TARGETED GENE DELIVERY.  
XX  
XX Hefeneider S, Merkins LS, Bennett RM, Seiss DC;  
XX  
XX WPI: 2002-188747/24.  
DR N-PSDB; AAD29592.  
XX  
XX New mammalian cell surface DNA receptor proteins and nucleic acids,  
PT useful for treating inflammation-associated diseases or conditions, e.g.  
PT otitis media, septic arthritis, or bacterial or viral infection that  
PT causes inflammation.  
XX  
XX Claim 2; Page 77-82; 94pp; English.  
XX  
XX The present invention relates to a nucleic acid comprising a sequence  
CC encoding a mammalian cell surface DNA receptor (DNA-R), a soluble  
CC mammalian cell surface DNA-R, or a DNA-binding fragment of a mammalian  
CC DNA-R. Mammalian DNA-R proteins may be used to screen compounds that  
CC effect DNA binding to cells in vivo and in vitro, and for treating  
CC inflammation-associated diseases and conditions including otitis media,  
CC septic arthritis, and any bacterial or viral infection that causes  
CC inflammation by interaction with the DNA-R. These may also be used to  
CC screen compounds that modulate binding, uptake and expression. The  
CC nucleic acid probes are useful for isolating mammalian species analogues,  
CC and for detecting mammalian DNA-R gene expression in cells and tissues.  
CC Recombinant expression constructs are useful in molecular biology to  
CC transform cells which do not ordinarily express a DNA-R, and the cells  
CC are useful as intermediates for making cell membrane preparations for  
CC receptor binding assays, which are subsequently useful in drug screening.  
CC The recombinant constructs are also useful in gene therapy. The present  
CC sequence is human DNA-R. DNA\_R gene is located on chromosome 9q34  
XX  
XX Sequence 1191 AA;  
SQ  
Query Match 100.0%; Score 6215; DB 5; Length 1191;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MPVQAQWTEFISCPICYNEDFNHVKPISLGCSTVCKTCLNKLHRAKCPDQTAINTD 60  
1 MPVQAQWTEFISCPICYNEDFNHVKPISLGCSTVCKTCLNKLHRAKCPDQTAINTD 60  
61 IDVLVFNFALLQVGAQVDPHQSIKLSNLGENKHVEVAKKVEDLALYLKPLSGGKGVAS 120  
61 IDVLVFNFALLQVGAQVDPHQSIKLSNLGENKHVEVAKKVEDLALYLKPLSGGKGVAS 120  
121 LNQSALSRPMQKLVTLVNCQVLEEGRVRAARAARSLGERTVTELILOHQNPOOLSANL 180  
121 LNQSALSRPMQKLVTLVNCQVLEEGRVRAARAARSLGERTVTELILOHQNPOOLSANL 180  
181 WAAVARGCQFLGPAMQEBEALKVLLALDGSALSRKVLVLFVQVORLEPRFPQAKSTIG 240  
181 WAAVARGCQFLGPAMQEBEALKVLLALDGSALSRKVLVLFVQVORLEPRFPQAKSTIG 240  
241 HVVQLLYRASCFTVKTRDESSLMQLKEEFRSYEARLREHDAQIVHIAEAGLRISPEOW 300  
241 HVVQLLYRASCFTVKTRDESSLMQLKEEFRSYEARLREHDAQIVHIAEAGLRISPEOW 300  
301 SSLLYCDLAHKSHMOSIIDKLOSPESPAKSVOELTIVLQRTGDPANLRLRPHLELLANI 360  
301 SSLLYCDLAHKSHMOSIIDKLOSPESPAKSVOELTIVLQRTGDPANLRLRPHLELLANI 360  
361 DPNPAVSTPWQLENAMAVKTVVHGLVDFQNTYSRKGHETFPQOPNSKYKTSMCRDLR 420  
361 DPNPAVSTPWQLENAMAVKTVVHGLVDFQNTYSRKGHETFPQOPNSKYKTSMCRDLR 420  
421 QGGCPRGNTCTFAHSQEBELEKYRLRNKKINATVTFPPLLNKVGNNVTVTAGNVISVI 480  
421 QGGCPRGNTCTFAHSQEBELEKYRLRNKKINATVTFPPLLNKVGNNVTVTAGNVISVI 480

QY 481 GSTETTKGIVPSTNGISNAENSVSQISRSTDSTIRALETVKVKGVKVGANGONAGPSAD 540  
DB 481 GSTETTKGIVPSTNGISNAENSVSQISRSTDSTIRALETVKVKGVKVGANGONAGPSAD 540  
QY 541 SVTENKIGSPKTPVSNVAATSAAGSNVGTSLNSVPQKSSPFLTRVPVYPHSENIQIFQ 600  
DB 541 SVTENKIGSPKTPVSNVAATSAAGSNVGTSLNSVPQKSSPFLTRVPVYPHSENIQIFQ 600  
QY 601 DPRTQIPPEVPOYQOTGYPPPTVPAGVPCVPRFVRSNNVPSSLPASMPYADHYST 660  
DB 601 DPRTQIPPEVPOYQOTGYPPPTVPAGVPCVPRFVRSNNVPSSLPASMPYADHYST 660  
QY 661 FSPDRMNSSPYQPPPOPGVPPVPSGMATPVVDSRRIRWRPMYQRDDIIRSNLSLPPM 720  
DB 661 FSPDRMNSSPYQPPPOPGVPPVPSGMATPVVDSRRIRWRPMYQRDDIIRSNLSLPPM 720  
QY 721 DVMSSVYQTSRLRERYNSLDGYYSVACQPPSEPRTTVPLPREPCGHLTKTSCBQIRRRPD 780  
DB 721 DVMSSVYQTSRLRERYNSLDGYYSVACQPPSEPRTTVPLPREPCGHLTKTSCBQIRRRPD 780  
QY 781 QMAQYHTOKAPLVSTLPAVATQSPPTPSPLSVDPADFSVSCTKFEEDHLSHYSPWS 840  
DB 781 QMAQYHTOKAPLVSTLPAVATQSPPTPSPLSVDPADFSVSCTKFEEDHLSHYSPWS 840  
QY 841 CGTIGSCINADISPEPKDVIANSNVLMDLSDGVKRRVHLFETQRTTKEEDPIIPFSDGP 900  
DB 841 CGTIGSCINADISPEPKDVIANSNVLMDLSDGVKRRVHLFETQRTTKEEDPIIPFSDGP 900  
QY 901 IISKWGAISRSTGTYHTTDPVQATASQGSATKPIVSVDYVYVNAVDSRWSSYNGEATS 960  
DB 901 IISKWGAISRSTGTYHTTDPVQATASQGSATKPIVSVDYVYVNAVDSRWSSYNGEATS 960  
QY 961 SAHYVERDRFVITDLSGHRKHSSTGDLISLELOAKSNLSLLQREANALAMQOKWNSLDE 1020  
DB 961 SAHYVERDRFVITDLSGHRKHSSTGDLISLELOAKSNLSLLQREANALAMQOKWNSLDE 1020  
QY 1021 GRHLTLNLLSKEIELRNGELQSDYTEDATDKPDRDIELELSALDTPDQSGSEPIEIL 1080  
DB 1021 GRHLTLNLLSKEIELRNGELQSDYTEDATDKPDRDIELELSALDTPDQSGSEPIEIL 1080  
QY 1081 DIQGISQNDQLNGMAVENGHPVQHQKBPPOKQKOSLGEDHVLBEQKTLIPVTSQF 1140  
DB 1081 DIQGISQNDQLNGMAVENGHPVQHQKBPPOKQKOSLGEDHVLBEQKTLIPVTSQF 1140  
QY 1141 SQPLPVSISNASCLPITTSVSGNLIILKTHVMSDKNDFLKPVANGKWVNS 1191  
DB 1141 SQPLPVSISNASCLPITTSVSGNLIILKTHVMSDKNDFLKPVANGKWVNS 1191  
RESULT 2  
AAE18609 standard; protein; 1191 AA.  
XX  
XX AAE18609;  
XX AC  
XX 17-MAY-2002 (first entry)  
XX  
XX Human cell surface DNA receptor (DNA-R) #2.  
DE Human; cell surface DNA receptor; DNA-R; inflammation-associated disease;  
KW otitis media; septic arthritis; drug screening; gene therapy; virucide;  
KW antiinflammatory; antibacterial; immunosuppressive; chromosome 9q34;  
KW auditory.  
XX  
XX Homo sapiens.  
XX  
XX WO200210392-A2.  
XX  
XX 07-FEB-2002.  
XX  
XX 01-AUG-2001; 2001WO-US024351.  
XX  
XX 01-AUG-2000; 2000US-0222624P.  
XX

XX (UYOR-) UNIV OREGON HEALTH SCI.  
PA (USGO ) US DEPT VETERANS AFFAIRS.  
PA (TARG-) TARGETED GENE DELIVERY.  
XX  
PI Hefeneider S, Merkins LS, Bennett RM, Seiss DC;  
XX  
DR WPI: 2002-188747/24.  
N-PSDB; AAD29593.  
XX  
PT New mammalian cell surface DNA receptor proteins and nucleic acids,  
PT useful for treating inflammation-associated diseases or conditions, e.g.  
PT otitis media, septic arthritis, or bacterial or viral infection that  
PT causes inflammation.  
XX  
PS Disclosure; Page 88-94; 94pp; English.  
XX  
CC The present invention relates to a nucleic acid comprising a sequence  
CC encoding a mammalian cell surface DNA receptor (DNA-R), a soluble  
CC mammalian cell surface DNA-R, or a DNA-binding fragment of a mammalian  
CC DNA-R. Mammalian DNA-R proteins may be used to screen compounds that  
CC effect DNA binding to cells in vivo and in vitro, and for treating  
CC inflammation-associated diseases and conditions including otitis media,  
CC septic arthritis, and any bacterial or viral infection that causes  
CC inflammation by interaction with the DNA-R. These may also be used to  
CC screen compounds that modulate binding, uptake and expression. The  
CC nucleic acid probes are useful for isolating mammalian species analogues,  
CC and for detecting mammalian DNA-R gene expression in cells and tissues.  
CC Recombinant expression constructs are useful in molecular biology to  
CC transform cells which do not ordinarily express a DNA-R, and the cells  
CC are useful as intermediates for making cell membrane preparations for  
CC receptor binding assays, which are subsequently useful in drug screening.  
CC The recombinant constructs are also useful in gene therapy. The present  
CC sequence is human DNA-R. DNA\_R gene is located on chromosome 9q34  
XX  
SQ Sequence 1191 AA;

Query Match 99.9%; Score 6207; DB 5; Length 1191;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPVQAQWTEFLSCPICYNFEDNVHKPISLGCSTVCKTCLNKLHKKACPFQDTAINTD 60  
DB 1 MAVQAQWTEFLSCPICYNFEDNVHKPISLGCSTVCKTCLNKLHKKACPFQDTAINTD 60

QY 61 IDVLPVNFALLQVGAQVDPHOSIKLSNKGKHVEVAKKCVEDLALYKPLSGKGKVAS 120  
DB 61 IDVLPVNFALLQVGAQVDPHOSIKLSNKGKHVEVAKKCVEDLALYKPLSGKGKVAS 120

QY 121 LQNSALSRPQWKLVTLVNQLVEEGRVRAMRAARSLGERTVTTELILQHONPOQLSANL 180  
DB 121 LQNSALSRPQWKLVTLVNQLVEEGRVRAMRAARSLGERTVTTELILQHONPOQLSANL 180

QY 181 WAAVRARGCQFLGPMQEEALKVLVLALEDGSSALSRKVLVLFVQRLPEPRFPOAKSTISG 240  
DB 181 WAAVRARGCQFLGPMQEEALKVLVLALEDGSSALSRKVLVLFVQRLPEPRFPOAKSTISG 240

QY 241 HVVQLLYRASCFFVKTRDESSLMQLKEFRSRYEARLRREHDAQIVHIAEAGLRISPEQW 300  
DB 241 HVVQLLYRASCFFVKTRDESSLMQLKEFRSRYEARLRREHDAQIVHIAEAGLRISPEQW 300

QY 301 SLLYGLAHKSHMQSIIDKLQSPESFAKSVQELTIVLQRTGDPANLRRLRPHLELLANI 360  
DB 301 SLLYGLAHKSHMQSIIDKLQSPESFAKSVQELTIVLQRTGDPANLRRLRPHLELLANI 360

QY 361 DPNPDVSPTEOLENAMAVKTVVHGLVDFIQNSRKGHETPQOPNSKYKTSKMCRLDR 420  
DB 361 DPNPDVSPTEOLENAMAVKTVVHGLVDFIQNSRKGHETPQOPNSKYKTSKMCRLDR 420

QY 421 QQGGCPRGNTCTFAHSEEELEKYLRLNKKINATVTFPPLNKKVGNVNTVTTAGNVISVI 480  
DB 421 QQGGCPRGNTCTFAHSEEELEKYLRLNKKINATVTFPPLNKKVGNVNTVTTAGNVISVI 480

QY 481 GSTETTGKIVPSTNGISNAENSVSOLISRSTDSITLRALETVKVKGVKGVANGQNAAGPSAD 540  
DB 481 GSTETTGKIVPSTNGISNAENSVSOLISRSTDSITLRALETVKVKGVKGVANGQNAAGPSAD 540

QY 541 SVTENKIGSPKTPVSNVAATSAAGPSNVGTENSVPOKSSPDLTPVVPVPHSENIQVFO 600  
DB 541 SVTENKIGSPKTPVSNVAATSAAGPSNVGTENSVPOKSSPDLTPVVPVPHSENIQVFO 600

QY 601 DPRTQIPPEVQYQOTGYPPPTVPAGVPCVPRFVRSNNVPSLPPASMPYADHYST 660  
DB 601 DPRTQIPPEVQYQOTGYPPPTVPAGVPCVPRFVRSNNVPSLPPASMPYADHYST 660

QY 661 FSPDRMNSSPYQPPPPQYGPVPGVPGMVAIVDSDRIWRPMPYQDDIIRNSLPM 720  
DB 661 FSPDRMNSSPYQPPPPQYGPVPGVPGMVAIVDSDRIWRPMPYQDDIIRNSLPM 720

QY 721 DVMHSSVQTSIRERYNSLDGYVSACOPPSPTTVPPLPREPCGHLTKTSCSEQIRRKPD 780  
DB 721 DVMHSSVQTSIRERYNSLDGYVSACOPPSPTTVPPLPREPCGHLTKTSCSEQIRRKPD 780

QY 781 QWAOYHTQKAPLVSSSTLPVATQSPPTPSPLFSVDFRADFSESVSGTKFEEDHLSHYSPWS 840  
DB 781 QWAOYHTQKAPLVSSSTLPVATQSPPTPSPLFSVDFRADFSESVSGTKFEEDHLSHYSPWS 840

QY 841 CQTIGSCINADSEPKDVIANSNVLMOLDGSDVKRRVHLFETQRTKEEDPIIPPSDGP 900  
DB 841 CQTIGSCINADSEPKDVIANSNVLMOLDGSDVKRRVHLFETQRTKEEDPIIPPSDGP 900

QY 901 IISKWGAISRSTGYHTTDPVQATASOGSATKPIVSVDYVYNVAVDSRWSSYGNKATS 960  
DB 901 IISKWGAISRSTGYHTTDPVQATASOGSATKPIVSVDYVYNVAVDSRWSSYGNKATS 960

QY 961 SAHYVERDRFIVTDLSGHRKHSSTGDLISLELQQAQKSNLILQREANALAMQKNSLDE 1020  
DB 961 SAHYVERDRFIVTDLSGHRKHSSTGDLISLELQQAQKSNLILQREANALAMQKNSLDE 1020

QY 1021 GRHLTLNLSKEIELRNGELQSDYTEDATDTKPRDRIELELSALDTPDQSGSEPIEIL 1080  
DB 1021 GRHLTLNLSKEIELRNGELQSDYTEDATDTKPRDRIELELSALDTPDQSGSEPIEIL 1080

QY 1081 DIQLGISSNDOLLNGMAVENGHVQHQKEPKKQKQSLGSDHVLLEQKTLIPVTSF 1140  
DB 1081 DIQLGISSNDOLLNGMAVENGHVQHQKEPKKQKQSLGSDHVLLEQKTLIPVTSF 1140

QY 1141 SLPPLVSIASCLPITTSVSAGNLILKTHVMSDKNDPLKPVANGKMYNS 1191  
DB 1141 SLPPLVSIASCLPITTSVSAGNLILKTHVMSDKNDPLKPVANGKMYNS 1191

RESULT 3  
AAE18613  
ID AAE18613 standard; protein; 1191 AA.  
XX  
AC AAE18613;  
XX  
DT 17-MAY-2002 (first entry)  
XX  
DE Human cell surface DNA receptor (DNA-R) #1 mutant, C431A.  
XX  
KW Human; cell surface DNA receptor; DNA-R; inflammation-associated disease;  
KW otitis media; septic arthritis; drug screening; gene therapy; virucide;  
KW antinflammatory; antibacterial; immunosuppressive; chromosome 9q34;  
XX  
XX auditory; mutant; mutein.  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT Misc-difference 431 /note= "Wild type Cys substituted with Ala"  
XX  
XX WO200210392-A2.  
XX  
XX 07-FEB-2002.

XX 01-AUG-2001; 2001WO-US024351.  
 XX 01-AUG-2000; 2000US-0222624P.  
 PR (UYOR-) UNIV OREGON HEALTH SCI.  
 PA (USGO ) US DEPT VETERANS AFFAIRS.  
 PA (TARG-) TARGETED GENE DELIVERY.  
 XX Hefeneider S, Merkins LS, Bennett RM, Seiss DC;  
 PI WPI; 2002-188747/24.  
 XX New mammalian cell surface DNA receptor proteins and nucleic acids,  
 PT useful for treating inflammation-associated diseases or conditions, e.g.  
 PT otitis media, septic arthritis, or bacterial or viral infection that  
 PT causes inflammation.  
 XX Example 2; Page; 94pp; English.  
 XX The present invention relates to a nucleic acid comprising a sequence  
 CC encoding a mammalian cell surface DNA receptor (DNA-R), a soluble  
 CC mammalian cell surface DNA-R, or a DNA-binding fragment of a mammalian  
 CC DNA-R. Mammalian DNA-R proteins may be used to screen compounds that  
 CC effect DNA binding to cells in vivo and in vitro, and for treating  
 CC inflammation-associated diseases and conditions including otitis media,  
 CC septic arthritis, and any bacterial or viral infection that causes  
 CC inflammation by interaction with the DNA-R. These may also be used to  
 CC screen compounds that modulate binding, uptake and expression. The  
 CC nucleic acid probes are useful for isolating mammalian species analogues,  
 CC and for detecting mammalian DNA-R gene expression in cells and tissues.  
 CC Recombinant expression constructs are useful in molecular biology to  
 CC transform cells which do not ordinarily express a DNA-R, and the cells  
 CC are useful as intermediates for making cell membrane preparations for  
 CC receptor binding assays, which are subsequently useful in drug screening.  
 CC The recombinant constructs are also useful in gene therapy. The present  
 CC sequence is human DNA-R mutant. DNA R gene is located on chromosome 9q34.  
 CC Note: The present sequence is not shown in the specification, but is  
 CC derived from the wild type DNA-R sequence shown in sequence listing (page  
 CC 77-82) of the specification  
 XX Sequence 1191 AA;  
 SQ  
 Query Match 99.9%; Score 6206; DB 5; Length 1191;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MPVQAQWTEFLSCPCYNEFDENVHVKPISLGCSTVCKTCLNKLHRRKACPPDQTAINTD 60  
 DB 1 MPVQAQWTEFLSCPCYNEFDENVHVKPISLGCSTVCKTCLNKLHRRKACPPDQTAINTD 60  
 QY 61 IDVLVPNFALLQVGAQVPDQHSIKLSNLGENKHVEAVKCVEDIALYLKPLSGKGVAS 120  
 DB 61 IDVLVPNFALLQVGAQVPDQHSIKLSNLGENKHVEAVKCVEDIALYLKPLSGKGVAS 120  
 QY 121 LNCALSRLPMORLVTLVNCQVLEEGVRVRAARSIGERTVTELILOHQNPOOLSANL 180  
 DB 121 LNCALSRLPMORLVTLVNCQVLEEGVRVRAARSIGERTVTELILOHQNPOOLSANL 180  
 QY 181 WAAVARGCQFGLPAMQAEALXVLLALEDDGALSRLKVLVLFVQRLERFPFQASKTSIG 240  
 DB 181 WAAVARGCQFGLPAMQAEALXVLLALEDDGALSRLKVLVLFVQRLERFPFQASKTSIG 240  
 QY 241 HVVQLLYRASCPCVTKRDESSLMQKBEFRSYEARLREHDAQIVHIAEAGLRISPCQW 300  
 DB 241 HVVQLLYRASCPCVTKRDESSLMQKBEFRSYEARLREHDAQIVHIAEAGLRISPCQW 300  
 QY 301 SLLLYGDLAHKSHMOSIIDKLOSPSPFAKSVQELTIVLQRTGDPANLRLPHELLANI 360  
 DB 301 SLLLYGDLAHKSHMOSIIDKLOSPSPFAKSVQELTIVLQRTGDPANLRLPHELLANI 360  
 QY 361 DPNDAVSPTWEOLENAMVAVKTVVHGLVDFIQNYSRKGHETPQPQPNKYKTSRCRDLR 420

DB DPNDAVSPTWEOLENAMVAVKTVVHGLVDFIQNYSRKGHETPQPQPNKYKTSRCRDLR 420  
 QY QGCGPRGTCTFAHSEEELEKYRLRNKINATVTRTPLLNKGVNNTVTTTAGNVISVI 480  
 DB QGCGPRGTCTFAHSEEELEKYRLRNKINATVTRTPLLNKGVNNTVTTTAGNVISVI 480  
 QY GSTETTKIVPSTNGISNAENSVSQLSRSRSTDLRALETVKYKGVKGANGQNAAGPSAD 540  
 DB GSTETTKIVPSTNGISNAENSVSQLSRSRSTDLRALETVKYKGVKGANGQNAAGPSAD 540  
 QY SVTENKIGSPPKTPVSNVAATSAGPSNVGTGLNSVPQKSSPFLTRVPVYPHSENIQIFQ 600  
 DB SVTENKIGSPPKTPVSNVAATSAGPSNVGTGLNSVPQKSSPFLTRVPVYPHSENIQIFQ 600  
 QY DPRQIPEVPOYQOTGYYPBPPTVPAGVAPCVPRFVRNNVPSSLPSPAMPADHYST 660  
 DB DPRQIPEVPOYQOTGYYPBPPTVPAGVAPCVPRFVRNNVPSSLPSPAMPADHYST 660  
 QY FSPDRMNSSPYQPPPPQYGPVPPVPSGMYPVYDSRRIRWRPPMYQRDDIIRSNLPPM 720  
 DB FSPDRMNSSPYQPPPPQYGPVPPVPSGMYPVYDSRRIRWRPPMYQRDDIIRSNLPPM 720  
 QY DVMSSVYQTSRLRRYNSLDGYYSVACOPPEPRTTVPDPCHLTKTSCQIRRRPD 780  
 DB DVMSSVYQTSRLRRYNSLDGYYSVACOPPEPRTTVPDPCHLTKTSCQIRRRPD 780  
 QY QMAQYHTQKAPLVGSTLPVATQSPTPSPPLSFVDFRADFSVSGTKFEEDHLSHSPWS 840  
 DB QMAQYHTQKAPLVGSTLPVATQSPTPSPPLSFVDFRADFSVSGTKFEEDHLSHSPWS 840  
 QY CGTIGSCINADISPPKDVIANSNVLMDLDSGDVKKRRVHLFETQRTTKEEDPIIPFSDGP 900  
 DB CGTIGSCINADISPPKDVIANSNVLMDLDSGDVKKRRVHLFETQRTTKEEDPIIPFSDGP 900  
 QY IISKWGAISRRTGYHTTDPVQATASQGSATKPIVSVDYVPYNAVDSRWSSYVNEATS 960  
 DB IISKWGAISRRTGYHTTDPVQATASQGSATKPIVSVDYVPYNAVDSRWSSYVNEATS 960  
 QY SAHVYERDRFTVTDLSGHRKHSSTGDLISLELQAKNSLLOREANALAMQOKWNSLDE 1020  
 DB SAHVYERDRFTVTDLSGHRKHSSTGDLISLELQAKNSLLOREANALAMQOKWNSLDE 1020  
 QY GRHLTLNLLSKEIELRNGELQSDYTEDATDKPRDIELELSALDTPDQSQSEPIEIL 1080  
 DB GRHLTLNLLSKEIELRNGELQSDYTEDATDKPRDIELELSALDTPDQSQSEPIEIL 1080  
 QY DIQLGISQNDQLLNGMAVENGHPVQHQKPPKQKOSLGEDHVLBEQKTLIPVTSQF 1140  
 DB DIQLGISQNDQLLNGMAVENGHPVQHQKPPKQKOSLGEDHVLBEQKTLIPVTSQF 1140  
 QY SLOPLVSTSNASCPIPTTSVSAGNLIILKTHVMSDEKNDFLKPVANGKMWNS 1191  
 DB SLOPLVSTSNASCPIPTTSVSAGNLIILKTHVMSDEKNDFLKPVANGKMWNS 1191  
 RESULT 4  
 AAE18612  
 ID AAE18612 standard; protein; 1191 AA.  
 XX  
 AC AAE18612;  
 XX  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Human cell surface DNA receptor (DNA-R) #1 mutant, C416A.  
 XX  
 KW Human; cell surface DNA receptor; DNA-R; inflammation-associated disease;  
 KW otitis media; septic arthritis; drug screening; gene therapy; virucide;  
 KW antiinflammatory; antibacterial; immunosuppressive; chromosome 9q34;  
 KW auditory; mutant; mutein.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers





KW antiinflammatory; antibacterial; immunosuppressive; chromosome 9q34;  
 KW auditory; mutant; mutein.  
 XX Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT Misc-difference 431  
 FT /note= "Wild type Cys substituted with Ser"  
 XX  
 PN WO200210392-A2.  
 XX  
 XX 07-FEB-2002.  
 XX  
 XX 01-AUG-2001; 2001WO-US024351.  
 XX  
 XX 01-AUG-2000; 2000US-0222624P.  
 XX  
 XX (UYOR-) UNIV OREGON HEALTH SCI.  
 PA (USGO) US DEPT VETERANS AFFAIRS.  
 PA (TARG-) TARGETED GENE DELIVERY.  
 XX  
 PI Hefeneider S, Merkins LS, Bennett RM, Seiss DC;  
 XX WPI; 2002-188747/24.  
 XX  
 XX New mammalian cell surface DNA receptor proteins and nucleic acids,  
 PT useful for treating inflammation-associated diseases or conditions, e.g.  
 PT otitis media, septic arthritis, or bacterial or viral infection that  
 PT causes inflammation.  
 XX  
 XX Example 2; Page; 94pp; English.  
 XX  
 XX The present invention relates to a nucleic acid comprising a sequence  
 CC encoding a mammalian cell surface DNA receptor (DNA-R), a soluble  
 CC mammalian cell surface DNA-R, or a DNA-binding fragment of a mammalian  
 CC DNA-R. Mammalian DNA-R proteins may be used to screen compounds that  
 CC effect DNA binding to cells in vivo and in vitro, and for treating  
 CC inflammation-associated diseases and conditions including otitis media,  
 CC septic arthritis, and any bacterial or viral infection that causes  
 CC inflammation by interaction with the DNA-R. These may also be used to  
 CC screen compounds that modulate binding, uptake and expression. The  
 CC nucleic acid probes are useful for isolating mammalian species analogues,  
 CC and for detecting mammalian DNA-R gene expression in cells and tissues.  
 CC Recombinant expression constructs are useful in molecular biology to  
 CC transform cells which do not ordinarily express a DNA-R, and the cells  
 CC are useful as intermediates for making cell membrane preparations for  
 CC receptor binding assays, which are subsequently useful in drug screening.  
 CC The recombinant constructs are also useful in gene therapy. The present  
 CC sequence is human DNA-R mutant. DNA R gene is located on chromosome 9q34.  
 CC Note: The present sequence is not shown in the specification, but is  
 CC derived from the wild type DNA-R sequence shown in sequence listing (page  
 CC 77-82) of the specification  
 XX  
 SQ Sequence 1191 AA;  
 Query Match 99.8%; Score 6205; DB 5; Length 1191;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MPVQAAQWTEFLSCPIYNEFDENVHKPISLGCSTVCKTCLNKLHRCACPFDQTAINTD 60  
 DB 1 MPVQAAQWTEFLSCPIYNEFDENVHKPISLGCSTVCKTCLNKLHRCACPFDQTAINTD 60  
 QY 61 IDVLPNFALLQVGAQVDPHQSIKLSNLGENKHVEAKKCVEDLALYLPKLSGGKGVAS 120  
 DB 61 IDVLPNFALLQVGAQVDPHQSIKLSNLGENKHVEAKKCVEDLALYLPKLSGGKGVAS 120  
 QY 121 LNSALSRRPMQRKLVTLVNCQVVEEGRVRAARSIGERTVTELIQHONPOOLSANL 180  
 DB 121 LNSALSRRPMQRKLVTLVNCQVVEEGRVRAARSIGERTVTELIQHONPOOLSANL 180  
 QY 181 WAAVRARGCOFLGPAMQOEALKLVLLALEDGSALSRRKVLVLFVVQRLRPFQASKTSIG 240

RESULT 6  
 AAE18610  
 ID AAE18610 standard; protein; 1191 AA.  
 XX  
 AC AAE18610;  
 XX

181 WAAVRARGCOFLGPAMQOEALKLVLLALEDGSALSRRKVLVLFVVQRLRPFQASKTSIG 240  
 QY  
 241 HVVQLLYRASCFTKVTKEDESSLMQLKEEPSEYALREHDAQIVHIAWEAGLRISPEQW 300  
 DB  
 241 HVVQLLYRASCFTKVTKEDESSLMQLKEEPSEYALREHDAQIVHIAWEAGLRISPEQW 300  
 QY  
 301 SLLYGLDLAKSHMQSIIDKLQSPESPAKSQVELTIVLQRTGDPANLRLRPHLELLANI 360  
 DB  
 301 SLLYGLDLAKSHMQSIIDKLQSPESPAKSQVELTIVLQRTGDPANLRLRPHLELLANI 360  
 QY  
 361 DPNPDVSPTWEOLENAWAVKTVVHGLVDFIQNSYRKGHETPOQPNKYKTSMCRDLR 420  
 DB  
 361 DPNPDVSPTWEOLENAWAVKTVVHGLVDFIQNSYRKGHETPOQPNKYKTSMCRDLR 420  
 QY  
 421 QGGCPRGTNCTFAHSQBELEKYRLNKKINATVTPPELLNKVGNNVTVTAGNIVSI 480  
 DB  
 421 QGGCPRGTNCTFAHSQBELEKYRLNKKINATVTPPELLNKVGNNVTVTAGNIVSI 480  
 QY  
 481 GSTTETTKIVPSTNGISNAENSVSQIISRSTDSLRALETVKYKVGKVGANGQNAAGPSAD 540  
 DB  
 481 GSTTETTKIVPSTNGISNAENSVSQIISRSTDSLRALETVKYKVGKVGANGQNAAGPSAD 540  
 QY  
 541 SVTENKIGSPKPTPVSNVAATSAGPSNVGTSLNSVPQKSSPFLTRVPVYPHSENIQYFQ 600  
 DB  
 541 SVTENKIGSPKPTPVSNVAATSAGPSNVGTSLNSVPQKSSPFLTRVPVYPHSENIQYFQ 600  
 QY  
 601 DPTQIPEVPOYQOTGYPPPTVPAGVAPCVPRFVRSNNVPSSLPSPAMPYADHYST 660  
 DB  
 601 DPTQIPEVPOYQOTGYPPPTVPAGVAPCVPRFVRSNNVPSSLPSPAMPYADHYST 660  
 QY  
 661 FSPDRMNSPYQPPPOPGVPVPPVPSGMVAPVVDSSRIWRPMPYQDDDIIRSNLSLPPM 720  
 DB  
 661 FSPDRMNSPYQPPPOPGVPVPPVPSGMVAPVVDSSRIWRPMPYQDDDIIRSNLSLPPM 720  
 QY  
 721 DVMHSSVYQTSRLRERYNSLDGYISVACOPPSPTTVPVLPREPCHLKTSCEEQIRRPD 780  
 DB  
 721 DVMHSSVYQTSRLRERYNSLDGYISVACOPPSPTTVPVLPREPCHLKTSCEEQIRRPD 780  
 QY  
 781 QWQYHTOKAPLVSTLVPVATQSTPSPPLSVDFRADFSVSGTKFEEHLSHYSPWS 840  
 DB  
 781 QWQYHTOKAPLVSTLVPVATQSTPSPPLSVDFRADFSVSGTKFEEHLSHYSPWS 840  
 QY  
 841 CGTIGSCINADISPKDVIANSNVMDLSDGDKVRRVHLFETQRTKEEDPIIPFSDGP 900  
 DB  
 841 CGTIGSCINADISPKDVIANSNVMDLSDGDKVRRVHLFETQRTKEEDPIIPFSDGP 900  
 QY  
 901 IISKWGAISRSTGYHTTDPVQATAGSGSATKPIISVSDYVPYVNAVDSRWSSSYGNEATS 960  
 DB  
 901 IISKWGAISRSTGYHTTDPVQATAGSGSATKPIISVSDYVPYVNAVDSRWSSSYGNEATS 960  
 QY  
 961 SAHYVERDRFIVTDLGHRKHSSTGDLISLELOAKSNLILLOREANALAMQOKWNSLDE 1020  
 DB  
 961 SAHYVERDRFIVTDLGHRKHSSTGDLISLELOAKSNLILLOREANALAMQOKWNSLDE 1020  
 QY  
 1021 GRHLTLNLSKEIERNGELOSVDYTEDATDKPRDIELELSALDTPDGDGSEPIEIL 1080  
 DB  
 1021 GRHLTLNLSKEIERNGELOSVDYTEDATDKPRDIELELSALDTPDGDGSEPIEIL 1080  
 QY  
 1081 DIQLGISQNDQLNGMAVENGHPVQHQKPPKQKQSLGDEHVDHILBEQKTLIPVTSQF 1140  
 DB  
 1081 DIQLGISQNDQLNGMAVENGHPVQHQKPPKQKQSLGDEHVDHILBEQKTLIPVTSQF 1140  
 QY  
 1141 SQPLPVSISNASCUPITTSVAGNLIILKTHVMSDKNDPLKPVANGKWNVS 1191  
 DB  
 1141 SQPLPVSISNASCUPITTSVAGNLIILKTHVMSDKNDPLKPVANGKWNVS 1191



RESULT 7  
AD80934  
ID ADB80934 standard; protein; 1048 AA.  
XX  
AC ADB80934;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE RING-SH complex related protein, SEQ ID No 8.  
XX  
KW RING-SH 3; Gag protein; Gag late domain; P13K; actin; myosin; Hsp60;  
KW Hsp70; Hsp90; STAM2A; STAM2B; VHS-UIM; GTPase; E2 enzyme; teg101;  
KW cullin; RING-SH; clathrin; virucide; vaccine; antiviral; retrovirus;  
KW rhabdovirus; filovirus.  
XX  
OS Unidentified.  
XX  
PN WO2003033646-A2.  
XX  
PD 24-APR-2003.  
XX  
PF 31-JUL-2002; 2002WO-US024589.  
XX  
PR 31-JUL-2001; 2001US-0308958P.  
XX  
PR 09-NOV-2001; 2001US-0345846P.  
XX  
PA (PROT-) PROTEOLOGICS INC.  
XX  
PI Greener T, Moskowitz H, Reiss Y, Alroy I;  
XX  
DR WPI: 2003-393509/37.  
DR N-PSDB; ADB80975.  
XX  
PT New isolated protein complex comprising a RING-SH 3 polypeptide and  
PT another polypeptide, useful for detecting cells infected with a virus,  
PT and for treating viral disorders caused by retroviruses, rhabdoviruses,  
PT or filoviruses.  
XX  
XX Disclosure; Fig 8; 176pp; English.  
XX  
PS The invention relates to a novel isolated protein complex comprising a  
CC RING-SH 3 polypeptide and a polypeptide selected from a Gag protein, a  
CC Gag late domain, P13K, actin, myosin, Hsp60, Hsp70, STAM1, STAM2A,  
CC STAM2B, VHS-UIM, a GTPase, an E2 enzyme, teg101, a cullin, RING-SH, and a  
CC clathrin. The novel protein complex has virucide activity and can be used  
CC to treat disorders as part of a vaccine. The protein complex and  
CC composition are useful for detecting cells infected with a virus, for  
CC identifying agents having antiviral activity, and for treating viral  
CC disorders caused by retroviruses, rhabdoviruses, or filoviruses. This  
CC sequence is a protein comprising the RING-SH complex of the invention.  
XX  
SQ Sequence 1048 AA;  
Query Match 88.2%; Score 5483; DB 7; Length 1048;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1048; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MPVQAQWTEFLSCPICYNEFDENVHKPISLGCSTHTVCKTCLNKLHRKACPFDDQAINTD 60  
DB 1 MPVQAQWTEFLSCPICYNEFDENVHKPISLGCSTHTVCKTCLNKLHRKACPFDDQAINTD 60  
QY 61 IDVLVFNFAQLQVGAQVDPHQSIKLSNLGENKHVEVAKKCVEDLALYLKPLSGGKGVA 120  
DB 61 IDVLVFNFAQLQVGAQVDPHQSIKLSNLGENKHVEVAKKCVEDLALYLKPLSGGKGVA 120  
QY 121 LNQSALSRPQKRLVTLVNCQVVEEGRVRAARSLGERTVTELLIHOHPQOOLSANL 180  
DB 121 LNQSALSRPQKRLVTLVNCQVVEEGRVRAARSLGERTVTELLIHOHPQOOLSANL 180  
QY 181 WAAVRARGCOFLGPAQOEALKLVLLALEDDGSAKSRKVLVLFVQRLPRFPQAKSTIG 240  
DB 181 WAAVRARGCOFLGPAQOEALKLVLLALEDDGSAKSRKVLVLFVQRLPRFPQAKSTIG 240

QY 241 HVVQLLYRASCFCVTKRDESSLMQLKEEPRSYEARLREHDAQIVHIAEAGLRISPEQW 300  
DB 241 HVVQLLYRASCFCVTKRDESSLMQLKEEPRSYEARLREHDAQIVHIAEAGLRISPEQW 300  
QY 301 SLLYGDIAHKSHMQSIIDKLQSPESFAKSVOELTIVLQRTGDPANLRLRPHLELLANI 360  
DB 301 SLLYGDIAHKSHMQSIIDKLQSPESFAKSVOELTIVLQRTGDPANLRLRPHLELLANI 360  
QY 361 DPNPDVSPTEWOLENAMVAVKTVVHGLVDFIQNSRKGHETPOPOPNKYKTSRCRDLR 420  
DB 361 DPNPDVSPTEWOLENAMVAVKTVVHGLVDFIQNSRKGHETPOPOPNKYKTSRCRDLR 420  
QY 421 QOQCGPRGTCTFAHSGOELEKYLRLNKKINATVTRPPLLNKVGNNVTVTAGNIVSI 480  
DB 421 QOQCGPRGTCTFAHSGOELEKYLRLNKKINATVTRPPLLNKVGNNVTVTAGNIVSI 480  
QY 481 GSTETTGKIVPSTNGISNAENSVSQISRSRSTDLRALETVKVKGVKGANGAAGPSAD 540  
DB 481 GSTETTGKIVPSTNGISNAENSVSQISRSRSTDLRALETVKVKGVKGANGAAGPSAD 540  
QY 541 SVTENKIGSPKTPVSNVAATSAQPSNVGTSLNSVPOKSSPFLTRVPPVPHSENIQYFQ 600  
DB 541 SVTENKIGSPKTPVSNVAATSAQPSNVGTSLNSVPOKSSPFLTRVPPVPHSENIQYFQ 600  
QY 601 DPRTQIPPEVPQYPTQGYPPPTVPAGVPCVPRFVRSNNVPSLPPASPYADHYST 660  
DB 601 DPRTQIPPEVPQYPTQGYPPPTVPAGVPCVPRFVRSNNVPSLPPASPYADHYST 660  
QY 661 FSPDRMNSSPYQPPPPQYGPVPPVPSGMVAPVYDSRRRIWRPMPYQRDDIIRSNLPPM 720  
DB 661 FSPDRMNSSPYQPPPPQYGPVPPVPSGMVAPVYDSRRRIWRPMPYQRDDIIRSNLPPM 720  
QY 721 DMHSSVYQTSIRBRYNSLDGYYSVACOPPSEPRTPVLPREPCGHLKTSCEEQIRRPD 780  
DB 721 DMHSSVYQTSIRBRYNSLDGYYSVACOPPSEPRTPVLPREPCGHLKTSCEEQIRRPD 780  
QY 781 QWAOVHTQKAPLVSSTLPVATQSPPTPPLFSDVPRADFSESVSGTKFEEHLSHSPWS 840  
DB 781 QWAOVHTQKAPLVSSTLPVATQSPPTPPLFSDVPRADFSESVSGTKFEEHLSHSPWS 840  
QY 841 CGTIGSCINAIDSEPKDVIANSNVLMOLDSDGVKRRVHLFETQRTTKEEDPIIPFSDGP 900  
DB 841 CGTIGSCINAIDSEPKDVIANSNVLMOLDSDGVKRRVHLFETQRTTKEEDPIIPFSDGP 900  
QY 901 IISKWGAISRSGRTGHTTDPVQATASQGSATKPIVSVDYVPYVNAVDSRWSSYGNAT 960  
DB 901 IISKWGAISRSGRTGHTTDPVQATASQGSATKPIVSVDYVPYVNAVDSRWSSYGNAT 960  
QY 961 SAHYVERDRFIVTDLSGHRKHSSTGDLISLELQAKSNLLIQRANALAMQOKWNSLDE 1020  
DB 961 SAHYVERDRFIVTDLSGHRKHSSTGDLISLELQAKSNLLIQRANALAMQOKWNSLDE 1020  
QY 1021 GRHLTLNLLSKEIELRNGELQSDYTEDA 1048  
DB 1021 GRHLTLNLLSKEIELRNGELQSDYTEDA 1048  
RESULT 8  
AAM40354  
ID AAM40354 standard; protein; 869 AA.  
XX  
AC AAM40354;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 3499.  
XX  
KW Human; neutropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.

XX OS Homo sapiens.  
XX PN WO200153312-A1.  
XX XX  
XX PD 26-JUL-2001.  
XX PF  
XX PP 26-DEC-2000; 2000WO-US034263.  
XX PR 23-DEC-1999; 99US-00471275.  
XX PR 21-JAN-2000; 2000US-00488725.  
XX PR 25-APR-2000; 2000US-00552317.  
XX PR 20-JUN-2000; 2000US-00598042.  
XX PR 19-JUL-2000; 2000US-00620312.  
XX PR 03-AUG-2000; 2000US-00653450.  
XX PR 14-SEP-2000; 2000US-00682191.  
XX PR 19-OCT-2000; 2000US-00693036.  
XX PR 29-NOV-2000; 2000US-00727344.  
XX PA (HYSE-) HYSEQ INC.  
XX XX  
XX XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
XX PI Zhou P, Goodrich R, Drmanac RT;  
XX XX  
XX DR WPI: 2001-442253/47.  
XX DR N-PSDB; AA159510.  
XX XX  
XX PT Novel nucleic acids and polypeptides, useful for treating disorders such  
XX PT as central nervous system injuries.  
XX PS Example 6; SEQ ID NO 3499; 10078pp; English.  
XX XX  
XX CC The invention relates to human nucleic acids (AA157798-AA161369) and the  
XX CC encoded polypeptides (AA138542-AA142213) with neurotropic,  
XX CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful  
XX CC in gene therapy. A composition containing a polypeptide or polynucleotide  
XX CC of the invention may be used to treat diseases of the peripheral nervous  
XX CC system, such as peripheral nervous injuries, peripheral neuropathy and  
XX CC localised neuropathies and central nervous system diseases, such as  
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX CC utilisation of the activities such as: Immune system suppression,  
XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
XX CC assays for receptor activity, arthritis and inflammation, leukaemias and  
XX CC C.N.S disorders. Note: The sequence data for this patent did not form  
XX CC part of the printed specification  
XX SQ Sequence 869 AA;

Query Match 71.1%; Score 4416; DB 4; Length 869;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 844; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 196 MQEALKLVLLALEDGSLSRKVLVLFVQRLPRFPQASKTSIGHVQVLLYRASCCKVT 255  
DB 1 MQEALKLVLLALEDGSLSRKVLVLFVQRLPRFPQASKTSIGHVQVLLYRASCCKVT 60

QY 256 KRDESSLMOLKEBFRSYEARREHDAQIVHIAEAGLRISPEQWSLLYGLDLAKSHMQ 315  
DB 61 KRDESSLMOLKEBFRSYEARREHDAQIVHIAEAGLRISPEQWSLLYGLDLAKSHMQ 120

QY 316 SIIDKLQSPESFAKSVQELTIVLQRTGDPANLRLPHLELLANIDPNDAVPTWEQLE 375  
DB 121 SIIDKLQSPESFAKSVQELTIVLQRTGDPANLRLPHLELLANIDPNDAVPTWEQLE 180

QY 376 NAMVAVKTVVHGLVDFONTYSRKGHETPPQPNKYKTSKCRDLROCGGCGTCTCTFAH 435  
DB 181 NAMVAVKTVVHGLVDFONTYSRKGHETPPQPNKYKTSKCRDLROCGGCGTCTCTFAH 240

QY 436 SQEELEKYLRLNKKINATVTFPLLNKVGNNVTVTAGNVISVIGSTETTGGKIVPSTNG 495

Db 241 SQEELEKYLRLNKKINATVTFPLLNKVGNNVTVTAGNVISVIGSTETTGGKIVPSTNG 300  
QY 496 ISNAENSVSQILSRSTDSTLRALETVKVKVGKVGANGQNAAGPSADSVTENKIGSPKTPV 555  
Db 301 ISNAENSVSQILSRSTDSTLRALETVKVKVGKVGANGQNAAGPSADSVTENKIGSPKTPV 360  
QY 556 SNVAATSAGPSNVGTGLNSVQKSPFLTRVPVPPPHSENQYFQDPRTQIPFEVPOVPO 615  
Db 361 SNVAATSAGPSNVGTGLNSVQKSPFLTRVPVPPPHSENQYFQDPRTQIPFEVPOVPO 420  
QY 616 TGYVPPPTVPAGVAPCPVFRVRSNNVPESLPPASMPYADHYSTFSPRDRMNSPYQPP 675  
Db 421 TGYVPPPTVPAGVAPCPVFRVRSNNVPESLPPASMPYADHYSTFSPRDRMNSPYQPP 480  
QY 676 PPQYGPVPPVPSGMYAPVYDSRRIWRPMTYQDDIIRSNLPPMDVMHSSVYQTSLRER 735  
Db 481 PPQYGPVPPVPSGMYAPVYDSRRIWRPMTYQDDIIRSNLPPMDVMHSSVYQTSLRER 540  
QY 736 YNSLDGYYSVACQPPSEPRTTVPLPRPCGHLKTSCEQIRKPKDQWQAYHTQKAPLVSS 795  
Db 541 YNSLDGYYSVACQPPSEPRTTVPLPRPCGHLKTSCEQIRKPKDQWQAYHTQKAPLVSS 600  
QY 796 TLPVATQSPPTPSPLFSVDFRADPSESVSCTKFEEDHLSHYSPWSCGTIGSCINAIDSP 855  
Db 601 TLPVATQSPPTPSPLFSVDFRADPSESVSCTKFEEDHLSHYSPWSCGTIGSCINAIDSP 660  
QY 856 KDVIANSNAVLMDLSDGDKRRVHLFETQRTKEEDPIIPFGSDGPIISKWGAISRRTG 915  
Db 661 KDVIANSNAVLMDLSDGDKRRVHLFETQRTKEEDPIIPFGSDGPIISKWGAISRRTG 720  
QY 916 YHTTDPVQATASQGSATKPIVSVDYVYVNAVDSRWSSYGNATSSAHVVERDRFVTDL 975  
Db 721 YHTTDPVQATASQGSATKPIVSVDYVYVNAVDSRWSSYGNATSSAHVVERDRFVTDL 780  
QY 976 SGHRKHSSTGDLISLELOAKSNSILLQREANALAMQKWNLSLDEGRHLTLNLSKEIEL 1035  
Db 781 SGHRKHSSTGDLISLELOAKSNSILLQREANALAMQKWNLSLDEGRHLTLNLSKEIEL 840  
QY 1036 RNGLQ 1041  
Db 841 RNGLQ 846  
RESULT 9  
AA40355  
ID AA40355 standard; protein; 812 AA.  
XX  
AC AA40355;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 3500.  
XX  
KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US034263.  
XX  
PR 23-DEC-1999; 99US-00471275.  
XX PR 21-JAN-2000; 2000US-00488725.  
XX PR 25-APR-2000; 2000US-00552317.  
XX PR 20-JUN-2000; 2000US-00598042.  
XX PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.  
PR 14-SEP-2000; 2000US-00662191.  
PR 19-OCT-2000; 2000US-00693036.  
PR 29-NOV-2000; 2000US-00727344.  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
PI Zhou P, Goodrich R, Drmanac RT;  
XX  
DR WPI; 2001-442253/47.  
DR N-PSDB; AAI59511.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders such  
PT as central nervous system injuries.  
XX  
XX Example 6; SEQ ID NO 3500; 10078pp; English.  
XX  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
CC encoded polypeptides (AAI38642-AAI42213) with neurotrophic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemia and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
XX part of the printed specification  
XX  
SQ Sequence 812 AA;

Query Match 65.4%; Score 4066.5; DB 4; Length 812;  
Best Local Similarity 93.0%; Pred. No. 1.3e-302;  
Matches 787; Conservative 2; Mismatches 0; Indels 57; Gaps 1;

QY 196 MOEALKLVLLALEDGSALSRKVLVLFVQRLPEPRFPQASKTSIGHVQLLYRASCFTVT 255  
DB 1 MOEALKLVLLALEDGSALSRKVLVLFVQRLPEPRFPQASKTSIGHVQLLYRASCFTVT 60  
QY 256 KRDESSLMLKEEPRSYEARREHDAQIVHIAEAGLRISPEQWSSLLYGLDAKSHMQ 315  
DB 61 KRDESSLMLKEEPRSYEARREHDAQIVHIAEAGLRISPEQWSSLLYGLDAKSHMQ 120  
QY 316 SIIDKLQSPESPAKSVQELTIVLQRTGDPANLRLRPHLELLANIDPNPDVSPTEQLE 375  
DB 121 SIIDKLQSPESPAKSVQELTIVLQRTGDPANLRLRPHLELLANIDPNPDVSPTEQLE 180  
QY 376 NAMAVKTVHGLVDFIQNSRKGHETPOPNKYKTSKMRDLRQGGCPRGTCTFAH 435  
DB 181 NAMAVKTVHGLVDFIQNSRKGHETPOPNKYKTSKMRDLRQGGCPRGTCTFAH 240  
QY 436 SQEELEKYLRLNKKINATVTRTPLLKNKGVNNVTITAGNVISVIGSTETTKIYVSTNG 495  
DB 241 SQEELEKYLRLNKKINATVTRTPLLKNKGVNNVTITAGNVISVIGSTETTKIYVSTNG 300  
QY 496 ISNAENSVQLISRSSTDSTLRALETVKVKYKGVGANGQNAAGSADSVTENKIGSPKTPV 555  
DB 301 ISNAENSVQLISRSSTDSTLRALETVKVKYKGVGANGQNAAGSADSVTENKIGSPKTPV 360  
QY 556 SNAVATSAGPSNVGTNELNSVPKQSPFLTRVPVPPHSENIQYFODPRTQIFEFVQPQPQ 615  
DB 361 SNAVATSAGPSNVGTNELNSVPKQSPFLTRVPVPPHSENIQYFODPRTQIFEFVQPQPQ 420  
QY 616 TGYPPPPPTVPAGVAPCPFRFVRNNVNPSSILPPASMPYADHYSTFSPPDRNNSPYQPP 675  
DB 421 TGYPPPPPTVPAGVAPCPFRFVRNNVNPSSILPPASMPYADHYSTFSPPDRNNSPYQPP 480

QY 676 PPQYGPVPPVPSGMYAPVYDSRRIRWPPMYQRRDDIIRSNLPPMDVMHSSVYQTSLRER 735  
DB 481 PPQYGPVPPVPSGMYAPVYDSRRIRWPPMYQRRDDIIRSNLPPMDVMHSSVYQTSLRER 540  
QY 736 YNSLDGYYSVACQPPSPBPTTVPPLRBPFCGHLKTSCEEQIRKRPDQWQAQYHTQKAPLVSS 795  
DB 541 YNSLDGYYSVACQPPSPBPTTVPPLRBPFCGHLKTSCEEQIRKRPDQWQAQYHTQKAPLVSS 566  
QY 796 TLPVATQSPPTPSPPLFSVDPRADPSESVSQTKFEDHLSHYSPWSCGTIGSCINAIIDSEP 855  
DB 567 -----FSESVSQTKFEDHLSHYSPWSCGTIGSCINAIIDSEP 603  
QY 856 KDVIANSNVLMDLSDGDKRRVHLFETQRTKEDPIIPFSDGPIISKWGAISRSSRTG 915  
DB 604 KDVIANSNVLMDLSDGDKRRVHLFETQRTKEDPIIPFSDGPIISKWGAISRSSRTG 663  
QY 916 YHTTDPVQATASQGSATKPIISVDPYVYVNAVDSRWSSYGNATSSAHYVERDRPIVTDL 975  
DB 664 YHTTDPVQATASQGSATKPIISVDPYVYVNAVDSRWSSYGNATSSAHYVERDRPIVTDL 723  
QY 976 SGHRKHSSTGDLISLELOAKSNLILLOREANALAMQKWNLSDEGRHLTLNLLSKETEL 1035  
DB 724 SGHRKHSSTGDLISLELOAKSNLILLOREANALAMQKWNLSDEGRHLTLNLLSKETEL 783  
QY 1036 RNGELO 1041  
DB 784 RNGEVK 789  
RESULT 10  
ADC37580  
ID ADC37580 standard; protein; 1125 AA.  
XX AC ADC37580;  
XX DT 18-DEC-2003 (first entry)  
XX DE Human nucleic acid associated protein, NAAP-47.  
XX KW Human; nucleic acid associated protein; NAAP; cytostatic;  
KW antiarteriosclerotic; anticoagulant; neurotropic; neuroprotective;  
KW cerebroprotective; anti-HIV; antiallergic; antiinflammatory;  
KW thymic; gene therapy; cell proliferative disorder; cancer;  
KW atherosclerosis; neurological disorder; epilepsy; Huntington's disease;  
KW stroke; immune disorder; inflammatory disorder; AIDS; allergy;  
KW developmental disorder; Hypothyroidism; Cushing's syndrome; infection.  
XX OS Homo sapiens.  
XX PN WO2003046151-A2.  
XX PD 05-JUN-2003.  
XX PF 26-NOV-2002; 2002WO-US038445.  
XX PR 27-NOV-2001; 2001US-0333925P.  
XX PR 07-DEC-2001; 2001US-0340477P.  
XX PR 14-DEC-2001; 2001US-0340362P.  
XX PR 18-DEC-2001; 2001US-0342002P.  
XX PA (INCY-) INCYTE GENOMICS INC.  
XX PI Baughn MR, Becha SD, Bhatia U, Blake JJ, Burford N, Burrill JD;  
PI Chang H, Chawla NK, Elliott VS, Emerling BM, Forsythe IU, Gandhi AR;  
PI Gietzen KJ, Gorvad AE, Griffin JA, Hafalia AJA, Jackson JL, Ho A;  
PI Ison CH, Jackson AA, Jiang X, Jin P, Kable AE, Khare R, Lal PG;  
PI Lee EA, Lee S, Lee SY, Li JK, Lu DAM, Ramkumar J, Richardson TW;  
PI Sprague WW, Swarnakar A, Tang YT, Warren BA, Xu Y, Yao MG, Yue H;  
PI Zheng W;  
XX WPI; 2003-513642/48.  
DR N-PSDB; ADC37640.  
XX

PT New human nucleic acid associated proteins (NAAP), useful for diagnosing,  
PT treating and preventing diseases or conditions associated with the  
PT aberrant NAAP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, or  
PT infections.

PS Claim 1; SEQ ID NO 47; 383pp; English.

XX The present invention relates to novel human nucleic acid associated  
CC protein (NAAP) (I: ADC37534-ADC37593) and their coding sequences (J). The  
CC NAAPs and their coding sequences are useful in diagnosing, treating and  
CC preventing diseases or conditions associated with the decreased  
CC expression or over expression of NAAP, such as cell proliferative (e.g.  
CC cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's  
CC disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and  
CC developmental (e.g. hypothyroidism, Cushing's syndrome) disorders, or  
CC infections. These are also useful in assessing the effects of exogenous  
CC compounds on the expression of nucleic acid and amino acid sequences of  
CC NAAP.

XX Sequence 1125 AA;

Query Match 42.9%; Score 2665.5; DB 7; Length 1125;

Best Local Similarity 51.9%; Pred. No. 9.1e-195;

Matches 609; Conservative 141; Mismatches 295; Indels 129; Gaps 38;

QY 1 MPVQAQWTEFLSCPTCYNEFQNVHVKPISLGCSTVCKTCLNKLHRAKCPDQTAINTD 60

DB 1 MPVQAQWTEFLSCPTCYNEFQNVHVKPISLGCSTVCKTCLNKLHRAKCPDQTAINTD 60

QY 61 IDVLPVNFALLQVGAQVDPHOSIKL-SNLGENKHYEAKVCVEDIALYLKPLSGKGVA 119

DB 61 IELLPVNSALLQVGAQVDPHOSIKL-SNLGENKHYEAKVCVEDIALYLKPLSGARGV 120

QY 120 --SLNQSALSRPQKLVTLVNCQVVEEGRVVRAMRAARSLGERTVTELILOHQPQOLS 177

DB 121 LNSTQSVLSRPQKLVTLVNCQVVEEGRVVRAMRAARSLGERTVTELILOHQPQOLS 180

QY 178 ANLMAVAVRGCOFLGPAQOEALKVLVLALEDGSLSRKLVLFVQVLEPRFPQASKT 237

DB 181 SNLMAVAVRGCOFLGPAQOEALKVLVLALEDGSLSRKLVLFVQVLEPRFPQASKT 240

QY 238 SIGHVVQLLYRASCFTVKTRDESSLQMLKEEPSYEAELRRHDAQIVHAMEAGLRISP 297

DB 241 SIGHVVQLLYRASCFTVKTRDESSLQMLKEEPSYEAELRRHDSQIVQIAMEAGLRISP 300

QY 298 EQWSSLLYGLAHKSHMQSIIDKLQSPSPAKSVQELTIVLQRTGDPANLRLRPHLELL 357

DB 301 DQWSSLLYGLAHKSHMQSIIDKLQSPSPAKSVQELTIVLQRTGDPANLRLRPHLELL 360

QY 358 ANIDPNPDVSPVTPMOLENAWAVTVVHGLVDFIONYSRKGHETPQOPNSKYKTSMCR 417

DB 361 ANIDPNPDVSPVTPMOLENAWAVTVVHGLVDFIONYSRKGHETPQOPNSKYKTSMCR 420

QY 418 DLKQCGCGRGCTCTFAHSELEKYLRLNKKINATVTFPLINKVGNNVTVTAGNVI 477

DB 421 DMKQCGCGRGCTCTFAHSELEKYLRLNKKINATVTFPLINKVGNNVTVTAGNVI 477

QY 478 SVTGSTTETTKIVPS--TNGISNAENSVSQILSRSTSTLRALETVKVKGVKANGQNAAG 536

DB 478 PDEGAVDLPSRKPALPANGVSTGNTVTLIPRGTDG---SYDSSLKPGKI----- 525

QY 537 PSADSVTENKIGSPKTPVSNVAATAGPSNVGTLSNVPQKSSPFLTPV----- 588

DB 526 ---DHLSSAGSPDLPOLLESVPKSIAPVNP-----HSIPRGPADLPMPVTKPLQMPV 578

QY 589 ----YPPHSENTQYFQDPR-TQIPPEVPOYQOTGYPPPPPTVPAGVAPCVPRFVRNNV 642

DB 579 RGSQPLPAQOTDV-YQDPRGAAPPPEPAPYQGMVYTPP-----QCVSEFVR---- 626

QY 643 PESSLPASMPYADHYSTFSPRDMNSSYPQPPPPQYGVPPVPSGMVAPVYDSRRIWR 702

DB 627 PPSAPAPAPPYLDHYPPYL-QERVNSQYGTQFQ-----YPI-----YPSHYDGRVYP 677

QY 703 PPMYQRDIIIRSNLPPMDVMHSSV--YOTSURERYNSLDGYSVACQP-----PS-----E 752

DB 678 APSYTRBIFRESPI-PIEIPPAAPVPSRERYQOIESYYPVAPHPTQIRPSYLRP 736

QY 753 PRTTVPLPREPCGHLKTSCEQIRKPKDQWQYHTQKAPLVSSITLPVATQSTPPSPPLFS 812

DB 737 PYSRLPPPPPPQ-----HPSLDELHRRRKEMIAQLBERK---VISPPPPFA-PSLTLP-PTPH 787

QY 813 VDFRADFSESYSCTKFEEDHLSHVSPWCGTIGSCINAIIDSEPKDVIANSNVLMDLDSG 872

DB 788 PEEFLDELKLVAG-KYKGNDSYQSPWSCDTIGYIGTKDAKPKDVAAAGSVEMNVEK 846

QY 873 DVKRRVHLFEQRTKKE--EDPIIPSDGPIISKWGAISRSRTGYHTTDPVQATASQGS 930

DB 847 GM--RDQRLDLQRRAAETSDDLIPFGDRPTVSRFGAISRTSKTIYQAGPMQAMAPOGA 904

QY 931 ATKPIVSVDYYPYVNAVDSRW--SSYG--NEATSSAHYVERDRFIVTDLSGHRK--HSST 984

DB 905 PTKSINISDISPY--GTHGGWGAISPHQNTASOGHFSERERISMSVASHGKPLPSAE 962

QY 985 GDLISLELQ-----AKNSLL-LQREANALAMQ-----KWSLDEGRHLTLNL-- 1028

DB 963 REQLRELEQLNHOISQOTQLRGLEANTLAGOSQPPPPPPKPMGMISSSEQLSLELHQ 1022

QY 1029 LSKETELANGELQSDYTDATDKPDRIELESALDDE--PDGQSEPIEILDILQIGI 1086

DB 1023 VREIEGKTRLSME-NQCSLDMK-----SKLNTSKQAENGQPEQNKVPAEDLTL 1072

QY 1087 SSONQDLNGMAVENGHVPVQHQKPEPPKQKQSL 1120

DB 1073 TFPSSD-----VPNGSALTQENISLLSNKTSLSL 1099

RESULT 11

ADF55448 ID ADF55448 standard; protein; 1109 AA.

XX ADF55448;

XX DT 12-FEB-2004 (first entry)

XX DE Human novel polypeptide #12.

XX KW human; brain disease; mental disorder.

XX OS Homo sapiens.

XX PN JP2003245081-A.

XX PD 02-SEP-2003.

XX PF 25-FEB-2002; 2002JP-00047501.

XX PR 25-FEB-2002; 2002JP-00047501.

XX PA (KAZU-) ZH KAZUSA DNA KENKYUSHO.

XX DR WPI; 2003-857164/80.

XX DR N-PSDB; ADF55399.

XX PT New DNA derived from KG-1 cDNA library and encoded polypeptide, useful

XX PS for treating mental disorders.

XX PS Claim 4; SEQ ID NO 12; 401pp; Japanese.

XX CC The invention relates to a DNA which encodes a novel polypeptide. A  
CC vector containing the DNA is useful as a reagent in estimation of  
CC standard substance. The antibody is useful for detecting the vector  
CC containing the DNA and for screening substances and compounds that  
CC interact specifically with the vector containing the DNA. The vector  
CC containing the DNA is used as a pharmaceutical, in treatment or as  
CC preventive agent with respect to disease e.g., brain diseases preferably  
CC mental disorders. The present sequence represents the amino acid sequence







CC allergies, neurological disorders (e.g., Alzheimer's disease,  
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin  
CC disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,  
CC cardiovascular disorders, angiogenic disorders, kidney disorders,  
CC gastrointestinal disorders, pregnancy-related disorders, endocrine  
CC disorders, and infections. The proteins can also be used to aid wound  
CC healing and epithelial cell proliferation, to prevent skin aging due to  
CC sunburn, to maintain organs before transplantation, for supporting cell  
CC culture of primary tissues, to regenerate tissues, to identify their  
CC cognate ligands or binding partners, and in chemotaxis, and can be used  
CC as a food additive or preservative to modify storage properties.  
CC Antibodies specific for a protein of the invention can be used in  
CC alleviating symptoms associated with the disorders mentioned above, and  
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked  
CC immunosorbent assay (ELISA). The present sequence represents a human  
CC secreted protein fragment referred to in the disclosure of the invention  
XX  
SQ Sequence 421 AA;  
  
Query Match 33.9%; Score 2106; DB 4; Length 421;  
Best Local Similarity 99.2%; Pred. No. 1.7e-152;  
Matches 395; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 644 ESSLPPASMPYADHYSTFSPDRMNSSPYQPPPPQYGPVPPVPSGMYAPVYDSRRIMRP 703  
DB 1 ESSLPPGSMYPADHYSTFSPDRMNSSPYQPPPPQYGPVPPVPSGMYAPVYDSRRIMRP 60  
  
QY 704 PMYQRDDIIRNSLPPMDVHSSVYQTSRERNSLDGYYSVACQPPSPRTTVPILPREP 763  
DB 61 PMYQRDDIIRNSLPPMDVHSSVYQTSRERNSLDGYYSVACQPPSPRTTVPILPREP 120  
  
QY 764 CGHLKTSCEQIRRKPDQWQAQHTQKAPLVSSSTLPVATQSPPTPPLFSVDFRADFSSEV 823  
DB 121 CGHLKTSCEQIRRKPDQWQAQHTQKAPLVSSSTLPVATQSPPTPPLFSVDFRADFSSEV 180  
  
QY 824 SGTKFEDHLSHYSPWSCGTIGCINADSEPKDVIANSNAVLMDLDSGDKRRVHLFFET 883  
DB 181 SGTKFEDHLSHYSPWSCGTIGCINADSEPKDVIANSNAVLMDLDSGDKRRVHLFFET 240  
  
QY 884 QRTKEEDPIIPSDGPIISKWGAISRSTGYHTTDPVQATASQGSATKPISVSDYVPY 943  
DB 241 QRTKEEDPIIPSDGPIISKWGAISRSTGYHTTDPVQATASQGSATKPISVSDYVPY 300  
  
QY 944 VNAVDSRWSSYGNATSSAHYVERDRPIVTDLSGHRKHSSTGDLLELQQAQKNSLLIQ 1003  
DB 301 VNAVDSRWSSYGNATSSAHYVERDRPIVTDLSGHRKHSSTGDLLELQQAQKNSLLIQ 360  
  
QY 1004 REANALAMQKWNLSDEGRHLLTNLLSKIELRNGELQ 1041  
DB 361 REANALAMQKWNLSDEGRHLLTNLLSKIELRNGEVK 398  
  
RESULT 13  
AAM42140  
ID AAM42140 standard; protein; 306 AA.  
XX  
AC AAM42140;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 7071.  
XX  
KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
XX

PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000WO-US034263.  
XX  
PR 23-DEC-1999; 99US-00471275.  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-0052317.  
PR 20-JUN-2000; 2000US-00598042.  
PR 19-JUL-2000; 2000US-00620312.  
PR 03-AUG-2000; 2000US-00653450.  
PR 14-SEP-2000; 2000US-00662191.  
PR 19-OCT-2000; 2000US-00693036.  
PR 29-NOV-2000; 2000US-00727344.  
  
(HYSE-) HYSEQ INC.  
  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QH;  
XX Zhou P, Goodrich R, Drmanac RT;  
  
DR WPI; 2001-442253/47.  
DR N-PSDB; AAI61296.  
  
XX Novel nucleic acids and polypeptides, useful for treating disorders such  
XX as central nervous system injuries.  
  
PS Example 2; SEQ ID NO 7071; 10078pp; English.  
  
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the  
XX encoded polypeptides (AAM38642-AAM42213) with nontropic, and the  
XX immunosuppressant and cytostatic activity. The polynucleotides are useful  
XX in gene therapy. A composition containing a polypeptide or polynucleotide  
XX of the invention may be used to treat diseases of the peripheral nervous  
XX system, such as peripheral nervous injuries, peripheral neuropathy and  
XX localised neuropathies and central nervous system diseases, such as  
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX utilisation of the activities such as: Immune system suppression,  
XX Activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic  
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
XX assays for receptor activity, arthritis and inflammation, leukaemia and  
XX C.N.S disorders. Note: The sequence data for this patent did not form  
XX part of the printed specification  
  
SQ Sequence 306 AA;  
  
Query Match 25.6%; Score 1593; DB 4; Length 306;  
Best Local Similarity 98.7%; Pred. No. 2.6e-113;  
Matches 293; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
QY 514 TLRALETVKVKGVKGANGQNAAGPSADSVTENKIGSPKPTPVSNVAATSGAPSNVGTENL 573  
DB 4 TLRALETAKVKGVKGANGQNAAGPSADSVTENKIGSPKPTPVSNVAATSGAPSNVGTENL 63  
  
QY 574 SVPQKSSPFLTRVPVYPHSENIQYFQDPRTOIPEFVPOYPQTGYPPPTVPAGVAPCV 633  
DB 64 SVPQKSSPFLTRVPVYPHSENIQYFQDPRTOIPEFVPOYPQTGYPPPTVPAGVAPCV 123  
  
QY 634 PRFVRSNNVPESLPPASMPYADHYSTFSPDRMNSSPYQPPPPQYGPVPPVPSGMYAP 693  
DB 124 PRFVRSNNVPESLPPASMPYADHYSTFSPDRMNSSPYQPPPPQYGPVPPVPSGMYAP 183  
  
QY 694 VYDSRRIMPPMYQRDDIIRNSLPPMDVHSSVYQTSRERNSLDGYYSVACQPPSEP 753  
DB 184 VYDSRRIMPPMYQRDDIIRNSLPPMDVHSSVYQTSRERNSLDGYYSVACQPPSEP 243  
  
QY 754 RTTVPLPREPCGHLKTSCEQIRRKPDQWQAQHTQKAPLVSSSTLPVATQSPPTPPEPL 810  
DB 244 RTTVPLPREPCGHLKTSCEQIRRKPDQWQAQHTQKAPLVSSSTLPVATQSPPTPPEPL 300  
  
RESULT 14  
AAM42141



Db 1 MPIQAFSWTDFLNCPICCNEFAASQRCVPSVLCGCHTICKLCLTTLNLRQCPDQTVIVSD 60  
 Qy 61 IDVLVNEFALLQV-----GAQVPDHQSIKLSNGENKHYEVAKCVEDALALYKPL 112  
 Db 61 IDNLPINHALLOLVKDSLELLEAPPSPVQKPEHL---KYQLGQRCIEELALHLKSF 117  
 Qy 113 SGGKGVASINQSALSRPMQKLVTLVNCQVBEGRVAMRAARSIGERTVTTELILQHON 172  
 Db 118 LNLG---NGNLLTRPMLKLVTLVNCQVBEGRVAMRAARSIGERTVTTELILQHON 173  
 Qy 173 POOLSANLWAAVRARCCQFLGAMQBEALKVLVLALEDGALSARKVLVLFVQRLBPRPP 232  
 Db 174 POOLSSNLWAAVTRTCQFLGAMQBEVLKVLVLALEDGALSARKVLVLFVQRLBPRPP 233  
 Qy 233 QASKTSIGHVVOLLYRASCFTKTRDESSLMQKEEFYSALREHDAQIVHIAEAG 292  
 Db 234 QASKTSIGHVVOLLYRASCFTKREASDSLMQKEEFRTYDALREHDAQIVQIATEAG 293  
 Qy 293 LRISPQWSSLLYGLLAHSHMQSIIDKLQSPESPAKSVOELTIVLQRTGDPANLRLRP 352  
 Db 294 LRAIAPQWSSLLYGLVHSHMQSIIDKLQTPSSPAQSVOELVIALQRTSDPAKLSLHH 353  
 Qy 353 HLELLANIDPNDAVPTWEQLENAMVAVTVVHGLVDFIQNYS-RK-----GHETPQ 404  
 Db 354 HLKYLANIDPCAE-VAP-NSVLAEALDAVRHSVVGVLVNFLOHGVKKAQDGISSGGSGGT 411  
 Qy 405 PQNSKYKTSMDCLRQCGCPRGTNCTFAHSGEELKYLRLNKKINATVTFPLLNKVG 464  
 Db 412 TNSNPKYKISLCRDLNVRCVPRGSSCTFAHSGEVEERYARNR----- 455  
 Qy 465 VNNVTVTAGNVISVIGSTETGKIVPSTNGISNAENSVSOLISRSTDTLRALETVKV 524  
 Db 456 ----- 455  
 Qy 525 GKVGANGQNAAGPSADSVTENKIGSPKTPVGNVAATSAGPSNVGTFLNSVP--QKSSPF 582  
 Db 456 -----GKHKTPLA-----LQPPPAVGVGAIKKPLGEOGPP 487  
 Qy 583 LTRVPVYPPHSENIQYFQDPRTQIPPEVPQYQOTGYPPPTVPAGVPCVPRFVRSNNV 642  
 Db 488 LGNMPPLPMSP-MHYMGSPR-----GYLDP-----SLGLSP-----GGGL 522  
 Qy 643 PESSLPPAS---MP--YADHYSTF---SPDRMNSSPYQPPPPQYGPVPPVPSGMYAPV 694  
 Db 523 PPSHHSPIITRLIVPSRYDSRPSGFGGTPR-----IPSPREYOANPAP----- 566  
 Qy 695 YDSRRIRWPPMYQRDDIIRSNLPPM---DVHSSVYQTSLRERYNSLDGYYSVACQPP- 750  
 Db 567 --TORNANPNFVSNSNLHKGVMPLASGGDVFLA-----NPWEQAVLAQQOHP 614  
 Qy 751 -----SEPTTVPLPREPCGHLKTSCEBQ-IRKPDQWQYHTQKAPLVS-STLPVATOS 803  
 Db 615 QHPQQQQPPSSKPNRPLSILPATADTSFEKKPPNSVSIDLDRYPEVNVDAVPLFRS 674  
 Qy 804 PTPPSPLFSDVDFRADSESVSGTKFEEDHLSHYSPWSCGTIGSCINAIIDSEPKDVIANSN 863  
 Db 675 -----NNHNNNNNNHNNHNSGSSLLFWN-NTGKDSANFVS-----DSLDDDD 720  
 Qy 864 AVLMDLDSGDVKKRVHLFETQRTKEEDPIIPFSDGPIISKWGAISRSRTG 915  
 Db 721 ASTFDVPTGSSMLSIY-----GPICPK-----SSTTG 747

Search completed: August 4, 2005, 01:36:14  
 Job time : 185 secs

This Page Blank (Cont.)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 01:26:33 ; Search time 26 Seconds  
(without alignments)  
4407.468 Million cell updates/sec

Title: US-10-619-992-2  
Perfect score: 6215  
Sequence: 1 MPVQAAQWTEFLSCPICYNE.....MSEDKNDFLKPANGKMWNS 1191

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: PIR1:.\*  
2: PIR2:.\*  
3: PIR3:.\*  
4: PIR4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 843   | 13.6        | 1048   | 2 T23764 | hypothetical prote  |
| 2          | 219.5 | 3.5         | 1186   | 2 T19050 | hypothetical prote  |
| 3          | 207.5 | 3.3         | 1198   | 2 T49726 | hypothetical prote  |
| 4          | 206   | 3.3         | 1611   | 2 T38236 | hypothetical prote  |
| 5          | 203   | 3.3         | 786    | 2 T01456 | extensin homolog F  |
| 6          | 202   | 3.3         | 760    | 2 T06291 | extensin homolog T  |
| 7          | 194.5 | 3.1         | 1279   | 2 T18312 | hypothetical prote  |
| 8          | 193.5 | 3.1         | 744    | 2 E86255 | hypothetical prote  |
| 9          | 192.5 | 3.1         | 944    | 2 T28734 | hypothetical prote  |
| 10         | 191   | 3.1         | 368    | 2 C29356 | hydroxyproline-ric  |
| 11         | 191   | 3.1         | 2160   | 2 T20241 | hypothetical prote  |
| 12         | 190.5 | 3.1         | 760    | 2 F86387 | probable PCO kinase |
| 13         | 190.5 | 3.1         | 842    | 2 S60402 | protein kinase CLA  |
| 14         | 188.5 | 3.0         | 1819   | 2 T32008 | hypothetical prote  |
| 15         | 183   | 2.9         | 280    | 2 T03236 | extensin precursor  |
| 16         | 181.5 | 2.9         | 530    | 2 T48217 | hypothetical prote  |
| 17         | 181.5 | 2.9         | 2282   | 2 T42717 | DNA-binding protei  |
| 18         | 181   | 2.9         | 1257   | 2 T00486 | serine/threonine-s  |
| 19         | 179   | 2.9         | 2562   | 2 T14266 | Xin protein - chic  |
| 20         | 179   | 2.9         | 2783   | 1 A41948 | alpha-fetoprotein   |
| 21         | 178   | 2.9         | 1013   | 2 T33470 | hypothetical prote  |
| 22         | 177.5 | 2.9         | 416    | 2 JU0465 | extensin precursor  |
| 23         | 177.5 | 2.9         | 620    | 2 S06733 | hydroxyproline-ric  |
| 24         | 177.5 | 2.9         | 1375   | 2 S48375 | hypothetical prote  |
| 25         | 177   | 2.8         | 1812   | 2 I49350 | breast/ovarian can  |
| 26         | 175.5 | 2.8         | 429    | 2 T06296 | extensin-like prot  |
| 27         | 174.5 | 2.8         | 951    | 2 T47617 | extensin-like prot  |
| 28         | 173   | 2.8         | 1262   | 2 T25168 | hypothetical prote  |
| 29         | 173   | 2.8         | 1794   | 2 T38459 | hypothetical diver  |

nascent polypeptid  
alpha-A-crystallin  
extensin - common  
hypothetical prote  
apoptosis associat  
translation initia  
hypothetical prote  
extensin class 1 p  
hypothetical prote  
hypothetical prote  
zinc finger protei  
extensin homolog T  
hypothetical prote  
hypothetical prote  
ascites sialoglyco  
extensin-like prot

## ALIGNMENTS

### RESULT 1

T23764  
hypothetical protein M142.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T23764; T27112

R:McMurray, A.

submitted to the EMBL Data Library, May 1996

A:Reference number: Z19795

A:Accession: T23764

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1048 <W14>

A:Cross-references: UNIPROT:O45962; EMBL:Z73428; PIDN:CAA97810.1; GSPDB:GN00021; CESP:M142.6

A:Experimental source: clone M142

R:Steward, C.

submitted to the EMBL Data Library, September 1997

A:Reference number: Z20313

A:Accession: T27112

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1048 <W12>

A:Cross-references: EMBL:Z99276; PIDN:CAB16481.1; GSPDB:GN00021; CESP:M142.6

A:Experimental source: clone Y52D3

C:Genetics:

A:Gene: CESP:M142.6

A:Map position: 3

A:Introns: 21/3; 122/3; 381/3; 579/3; 647/2; 683/1; 783/2; 855/1; 871/2; 909/1; 993/1

Query Match 13.6%; Score 843; DB 2; Length 1048;

Best Local Similarity 30.6%; Pred. No. 9.8e-39;

Matches 238; Conservative 99; Mismatches 278; Indels 162; Gaps 23;

QY 4 QAAQWTEFLSCPICYNEFDENVHKPISGCSHTVCKTCLNKLHR--KACPFQDTAINTDI 61

Db 6 QGGQWQEVLCSSICNRHENE-TFLPVSLICGHVICRKAEPENQTKPCPHDDWTKTTHSP 64

QY 62 DVLPNFALLQVGAQVDPHQSIKLSNLGNKHCHYVAKKCVEDLALYLPKLSGKGVASL 121

Db 65 SEYPNNVALLSVI---FPRKQCMTSLGAVSEAKRVDQLSIQ-IAKFFREADSERG-GTV 119

QY 122 NQSALSRPMOKLVTLVNCQVLEEGRVRAAARSLSGERTVTETLLOHNPQOLSANLW 181

Db 120 SREISRTLQKRVALLCVQREVQGLTKLQKCGISERVMIETLSQSNTHVSSQLW 179

QY 182 AAVRARGCOFLGPAMQEEALKVLVLALEDGSAKSLRVLYLVVVQRLERPRPQASKTSIGH 241

Db 180 SAVRARGCOFLGPAMQDDVLRLILMTLETGECIARKNLVMYVVQTLASDPQVSKTCVGH 239

QY 242 VVQLYRASCFCVKVTRDESSLMQLKEEPRSYEARLRREDAQIVHVTAMAGLRISPEQWS 301

Db 240 VVQLLYRASCFCVNLKRDGESSLMQLKEEPRSYEARLRREDAQIVHVTAMAGLRISPEQWS 299

```
QY 302 SLLYGLAHKSHMOSIIDKIQSPESFAKSVQELTIVLQRTGDPAN-----LNRLRP 352
Db 300 ALLYADQSHRSQMOSIIDKIQSKNSIQQGVESLRAI-----AGQSMVLPAYRYFTQVIP 356
QY 353 HLELLANIDPNDAVSPWTEQLENAWAVKTVVHGLVDFIQNY-----SRKGHETPOQPNRS 409
Db 357 CLEFFAGIEHEDTSMRMIGDALHQIRILLK--LHCSQDDLKMPKEERGVILQAEVPGG 414
QY 410 KYKTSMCRDLROGGCPRTNCTFAHSQBELEKYRLRNKKINATVTPPLLNKVGNNTV 469
Db 415 -----MGCGPGGSG-----GAEAGRI----- 430
QY 470 TTTAGNVISVIGSTETTKIVPTNGCISNAENSVSQILSRSTDSTLRALETVKYKVGKA 529
Db 431 -----GGLHPLYSQIDETGHSISRTPKNSHNSPQ----- 461
QY 530 NGQNAAGPSADSVTENKIGSPKTPVSNVAATSAGSPNVGTSLNSVPQKSSPFLTRVPVY 589
Db 462 -----TPPKQPROKEYQNGIIPNRMG-----YSSDAPFFIFSHQOQ 497
QY 590 PHSENIQVQPRQIPEVQYQGTGYPPPTVPAGVPCVPRFVRSNNVPSSLPP 649
Db 498 PPP-----QPFNS--QHLPQRFGRGRGAGAPPPPPQP-----MPMLIGYD-----MPG 539
QY 650 ASMPYADHYSTSPDRMNSP-----YQPPPPQVGPVPPVPSGMYAPVYDSRRIWRPP 704
Db 540 APMQQATEVLT--ADGOMVNGTQFQVVIMQSPFHLPGGPVWMLPQQOMVPPPOSMTMPVGGP 598
QY 705 MYQRDDIIRSNLPPMDVMHSSV--YQTSLRERYNSLDGYYSV--ACQPPSPFRATTV 757
Db 599 M-----GPMGPMTPIQVQVPNTWMTATSPGSIYIPAAASPPGPPHTI 643

RESULT 2
T19050
hypothetical protein C07E3.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T19050
R:Matthews, P.
submitted to the EMBL Data Library, June 1995
A:Reference number: Z19066
A:Accession: T19050
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1186 <WIL>
A:Cross-references: UNIPROT:Q17786; EMBL:Z49908; PIDN:CAA90094.1; GSPDB:GN00020; CESP:CO
A:Experimental source: clone C07E3
C:Genetics:
A:Gene: CESP:C07E3.3
A:Map position: 2
A:Introns: 36/3; 79/3; 1128/3; 1160/3

Query Match 3.5%; Score 219.5; DB 2; Length 1186;
Best Local Similarity 19.1%; Pred. No. 0.00043;
Matches 215; Conservative 153; Mismatches 421; Indels 337; Gaps 49;

QY 265 QLKEFRSFEALR-REHDAQIVHIAEAGLRISPEQWSSLLYGLDLAKSHMQ--SIID-- 319
Db 59 QLRKREVELETIRIQHESREFIIDAYRQNPAPQ--GNLYNEFNHYDSMKHASPSPDPA 115
QY 320 -----KIQSPESFAKSVQELTI-----VLQRTGDPANLN----- 348
Db 116 TYFALKPMTEVNEGROFSIPTSTRYDQVPIDPSPELQRRVPSKQSPNYPKRSSVFG 175
QY 349 -RLRPHLELLANIDPNP-----DAVSPTWEQLENAWAVKTVVHGLVDFIQYSRKGHET 402
Db 176 GRVSP-LPQSEVDHPMPRYIPDMSPRYERTPKAQYCKTTV-----TYYTYG-DT 226
QY 403 POPQPNISKY-KTSMCRD--LRQGGCGPRT-----NCTFAHSQ----- 438
Db 227 PEPPRSTTYSKASKLETYSVKPYATYSRGTKLPKYDEVPRNTWYSRASEASKTSRTSLA 286
```

```
QY 439 -----ELEKYRL--RNKKINATVTPPLLNKVGNNTVTTTACNVISVIGSTETTKI 489
Db 287 STCSRASKLPKVDQIPRSPVSHLYCRIPTYDEVVQKSPRSITPTPMSHRVRIQTSPP 346
QY 490 VPSTNG-----ISNAENSVSQILSRSTDSTLRALETV-----KKVGKVGANGONAA 535
Db 347 PPSNSNTYVIPKVDNRSDSPRTVIIPFSSNSTTALDKVPQBTPIYRISK----- 397
QY 536 GPSADSVTENKIGSP-----KTPVSNVAATSAGSPNVGTSLNSVPQ-----KSP 581
Db 398 -PSQTKPTNSRVSSIFLSQSYHSMKDIWTMTSYRTTA--TNLKKQEPVIPRPTYRYSKTP 455
QY 582 FLTRPVVPPHSENIQY--FODPRTQIPPEVQYQGTGYPPPT-----V 625
Db 456 LQKEIPVPLERSVIRQPEYSDPGDSRIIVPFNSQNKYSPPPSRSTPIYHTLQRDV 515
QY 626 PAGVAPCVPRFVRSNNVPSSLPASMPYADHYSTSPDRMNSSPYQPPPPQYGPVPP 685
Db 516 PTPPLPRRPLYSR-----PPTISTASSVE--YIPRAPVPRRPLDYRPPDSTISPPRS 565
QY 686 VPSCMY-----APVDSRRRIWRPMPYQRDDIIRSNLPPMDVMHSSVYQTSLRER 735
Db 566 PPSSELYLTLEQEVQRAPL--SRR-----PLYSRPESTASAPRSP---QPKIYLTLOKE 615
QY 736 YNSLDGYYSVA-----CQPPSEPRRTVPLPR--EPCGH--LKTSCBEQIRRKPDQWAO 784
Db 616 VPHVPLYSCVAKPRPVYSTPPEEP--IVIPRANVPVSSITLYLTRPEVPGTFLQHY 672
QY 785 YHTQKAPLVS-----STLPVATQSPTPPSPFLFSVDFRADPSE--SVSGTKFE----- 829
Db 673 YRSRSSTYSVPPVEEASTRTVASPSRYTPPPSSSPKXHLTHLEREVPGTPLNHYISRR 732
QY 830 -----EDH-----LSHYSFWSGCTIGSCINATDSEPK----- 856
Db 733 STPIYSTTPEDRNVVSPSRYTPPPSVTMAAKLHLTLQPEVPESTPLQHYISRSASSVYN 792
QY 857 --DVITANSNAVIMDLSDGDKVRKRVHLFTQRTKEEDPIIPFSDGPIISKWGAISRST 914
Db 793 TPSVDSTSTRSLVSPSSVMPKXHLH----TLQQDVWTPPK-----VSRPST 838
QY 915 GYHTTDPVQATASQGSAT-----KPIVSVDYVPYVNAVDSRWSYGNATSSAHY 964
Db 839 PIYSTPPELEKPISTRITVTPSRRTTPQKLHQISLDQIP-----GTPLOHY 883
QY 965 VERDRPIVT---DLGHRKHSSTGDLISLE----- 991
Db 884 IYRSRSTYSIPLSDYQPSPTSPSSVIMNDDSPYEEKPKSYLGYTPKYTYTRIMPERPPK 943
QY 992 -----LQQAQNSILLQREANALAMQKWSLDEGRHLTLNLLSKET--ELRNGE 1039
Db 944 IISPEHPVTVSKSKSTQTLQDQ-----VVLQPKTHLYPEISYRRRRPISIHVPAESKST 1000
QY 1040 LOSDYTEDATDKPRDRIEELSALDTDEPQSQSEPIE--EILDIQIGISSQNDQLNGMA 1098
Db 1001 IISIYSNPIDDVPPTHVSELHAEKRMIVGIDDELQPEKNNTVYTVATAAIDNELLHRHP 1060
QY 1099 VENGHVPV-----QQHQKEPPPKQKQSLGEDHVLIEEQKTLIPVTS 1139
Db 1061 EVGRHPVTVLPAQTSYSGASAKRRSD-----DEAKNVLPQNC 1098

RESULT 3
T49726
hypothetical protein B23J21.390 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T49726
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49726
A>Status: preliminary
```







239 IGHVVQLLYRASCYKVTCKRDESDSLMQLKEBFRSYEARLRREHDAQIVHIANEAGLRISPE 298  
 336 HPWVHLHMHRA-----DLRLEERLEEQRNNMATKIQTWK-MERAMQ----- 437  
 299 QWSSLLYGDLAHSHMQSIIDKLOSPESFAKSVQELTIVLQRTGD-PANLRLRPHLELL 357  
 438 -----LEYTLGLPHVRRHQRI-----PRKY-RAMPDLDFREYVGDMPAD-----ELL 478  
 358 ANIDPNDAV--SPTWEQLENAMVAVKTVHGLVDFIQNYRKGHETPQ-----POP 407  
 479 G--DDADGVGGPGD-----HG-----HYGQQAEPQWTEVPSPAPVP 517  
 408 NSKYKTSMDRLRQOQCGPRGTCTFAHSQBELEKYRLNKKINAT--VRTFPLLNKVG 465  
 518 ESRHTG--DPRRPAQTP-----SARQSRKPPY-----NATALVRVAP-NEHAE 558  
 466 NNTVTTTGNVSVIGSTETTCKIVPSTNGISNAENSVLSISRSTDLRALETVKVG 525  
 559 RGDFTTLVPNATVVGTTQDT-----QTDTDVHHTYEG-DVLTVRTVTTKTITBDL----- 609  
 526 KVGANGQNAAGPSADSVTENKIGSPKTPVSNVAATSAGPSNVGTLSNVPKGS----- 580  
 610 -----VDSAGDDEESESDEGQ-----AAPLHVGA-----PPFORQOQOQPR 648  
 581 -----PFLTRVPVYPHSENIQYQDPR-----TQIPFEVQYQYPTGYPPTV----- 625  
 649 HATHSTHPPQLEPAP-YPPDDDADEYDEERECSWAERAPTELQOQOQAPQAPPPMRRE 707  
 626 -PAGVAPC-----VRFVRSNNV-----PSSSLP-----ASMPY----- 654  
 708 KQTQVHPCHVVTTQTRRAPRVNKTQEAAMPLEREEEAPRTSRQAATTASHFVRRET 767  
 655 ADHYSTFSPDRMNSSPQPPQYGP-VPPVPSGMVAPYVDSRRIRWPPMYOR----- 708  
 768 ATHN--PDRGAHRAHAPQPPPPAPIPORQPPVPSPL-----PSCRRRAGS 813  
 709 ---DDIIRSNLPPMDVHSSVYQTSLRERNVSLDGYTSVACQPSPEPRTTVPLPREPCG 765  
 814 RPAGGLEADNALYPTDSTPSIVARAP-ESRENSLP-----ERPPKADRPRVHL----- 861  
 766 HLKTSCEBQIRKKPDQMA-----QYHTQKAPLVST-----LPV----- 799  
 862 HNSRAENALDDEPSRSQARGALPQORHTQTAPPAHRSAGALEAPLDDYEDGNAPYQBEA 921  
 800 -----ATOSPFPPLFSVDFRAD-----PSESYS-GTKFREEDH-----LSH 835  
 922 RSEVIAGSDAQSSAPSRLXYVPDVRHNSRRELFRDLVEPNTAFQEEWNGCPAFLPLIGH 981  
 836 YSPWSCGTIGSCINAIDSEPKDVIANSNVAMLDLSDGDVKKRVHLFETQRTKEED----- 891  
 982 RRGSGRSSQSSMPTSSSESGOLVQHALGHNGSHPSKQKVPIQLYD-RRHQSGSDGQO 1040  
 892 ----PI----IPSDGPIISKWGA-----ISRSRTGYHTTDPVQA-----TASQ 928  
 1041 QSGMPEWYLTPTTNEYPLALPQAAQOPRTVKTTRDYKVIKTRVRVYTYTREPSSIMDSENT 1100  
 929 GSATPISVSDVPYVNAVDSWSSYGNATSSAH--YVERDRFIVTDLGSRKHSSTGD 986  
 1101 SSHSGSVSIGE-----SMSSAHWEVDSASVSGSTHVEYVDEHDAASVNSKPSN--- 1152  
 987 LLSLELQQAQKNS 999  
 1153 -VSAPRQOQOQPS 1164  
 RESULT 8  
 E86255  
 hypothetical protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C:Accession: E86255  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: E86255  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-744 <STO>  
 A:Cross-references: UNIPROT:065375; GB:AE005172; NID:g3157926; PIDN:AAC17609.1; GSPDB:G  
 C:Genetics:  
 A:Map position: 1  
 Query Match 3.1%; Score 193.5; DB 2; Length 744;  
 Best Local Similarity 23.2%; Pred. No. 0.0059;  
 Matches 116; Conservative 51; Mismatches 185; Indels 147; Gaps 25;  
 QY 394 NYSRKGHETPOQNSKYKTSMDRLRQOQCGPRGTCTFAHSQBELEKYRLNKKINA- 452  
 DB 202 NNNRFRGIPKMGNSPVSALVLAD-NNLGSCIPG---SIGOMKTLNELILSNDNLGTC 257  
 QY 453 -----TVRTFPLLNKGVNNVTVTAGNVISV-----IGSTETTCKIVPSTNGISNA 499  
 DB 258 LPQIGNLKVTYVDITSN-RLQGPLSPSVGNMKSLEELHVANNNAFTGVIPPSICQLSNL 316  
 QY 500 EN-----SVQLISRSTDSLRALETVKVKVGVANGQNAAGPSADSVTE 544  
 DB 317 ENFTYSSNYFSRPPICAAASLLADIVVNGTWNCTGLAR----QRSDKQCSLLARPVDC 372  
 QY 545 NKIG-----SPPKTPVSNVAATSAGPSNVGTLSNV 575  
 DB 373 SKFGCYNIFSPPTFKMSPEVRTLPPIYVYSPPPPSKMSFTVAYS-----PP 425  
 QY 576 PQKSSPFLTRVPVY-----PPHSE---NIQYFQDPRTOIPFEVQYQYPTGYPPTVPAG 628  
 DB 426 PPSKMSPSVRAYSPPPPYKMSPSVRAYPPPPSPSPPPPYVYSS--PPPPVYVSS 483  
 QY 629 VAPCVPRVRNNVPE-----SSLPPASMPYADHVTSTFPRDRMNSSPYQPPPPQYGPV- 683  
 DB 484 PPP--PPYVYSSPPPPPYVYSSPPP--PYV--YSS-PPPPYVYSSP-PPPPSPPPPCP 534  
 QY 684 ---PPVPSGMVAPYVDSRRIRWPPMYQRDDIIRSNLPPMDVHSSVYQTSLRERNVSLD 740  
 DB 535 ESSPPPPPVVYAPVTQS-----PP-----PPSPVYVYPTVQSP----- 567  
 QY 741 GYYSVACOPPS-----EPRTTVPLPREPCGHKLTSCCEQIRRKPDQMAQYHTQKAPLVST 796  
 DB 568 -----PPSPSVYVYPTVNSPPSPSVYVYPPVTVS-----PPPPSPVYVYPTVSPSPPP 615  
 QY 797 LPV-----ATQSGTPPSPLF 811  
 DB 616 SPLVYVYPTVSPSPSPSVY 634  
 RESULT 9  
 T28734  
 hypothetical protein F26G5.9 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T28734  
 R:Sammons, L.; Wohldmann, P.; Beck, C.  
 submitted to the EMBL Data Library, September 1997  
 A:Description: The sequence of C. elegans cosmid F26G5.  
 A:Reference number: Z30516  
 A:Accession: T28734  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-944 <SAM>  
 A:Cross-references: UNIPROT:016936; EMBL:AF022974; PIDN:AAC48041.1; GSPDB:GN00023; CBSP

A:Experimental source: strain Bristol N2; clone F26G5  
C:Genetics;  
A:Gene: CESP:F26G5.9  
A:Map position: 5  
A:Introns: 12/3; 48/3; 258/3; 327/3; 643/1; 916/2

```
Query Match      3.1%; Score 192.5; DB 2; Length 944;
Best Local Similarity 17.7%; Pred. No. 0.0096;
Matches 215; Conservative 146; Mismatches 410; Indels 445; Gaps 56;

QY 7 QWTEFLSPICYNEDENVHKPISLGCSTGCTKCLNKLHKACPFQDTAINTDIDLVPV 66
DB 14 QWAVVECPICNYID----KPMQMGCGHTLSTCIGRL-----VDQVKT 54
QY 67 NFALLQLVGAQVPDQHSIKLSNLGENKHYEVAKKVEDLALYLKPLSGKGVASLNQSA- 125
DB 55 NNA-----EELGRPR-----GPLLHPDDIGPGFGLGGFAD 85
QY 126 -LSRPMQKLVTLVNCQLVEEGRVRAEARSIGERTVTLELLOHNPQQLSANLWAAV 184
DB 86 QEDQHFDLPLRPRMPNDPNVRGI-----VLPFRGPE----- 122
QY 185 RARGQOFLGPAMQEBALKLVLLALEDDGALSRLVLVFLVYVORLEPRFPQASKTSIHVVQ 244
DB 123 -ARATEIKCEPKRKT-----LVPADGLPVNVRVQE--IVQKVAFLPKDR-----HLVK 168
QY 245 LLYRASCFTVTRDESSLMQLKEBSRSEALRREHDAQIVHIAEAGLRISPEOWSSLL 304
DB 169 LCNQ--CEAVLSQGVYDFCSQCEETGRK-----ICSTCAIRL----- 203
QY 305 YGDLAKSHMQSIIDKLQSPESFAKSVQELTIVLORTGDPANLRLRPHLELLANIDNP 364
DB 204 -----HNGH-----QVKEKALTSDVREMKKISDASH-----AAFQSLNKLKRL 245
QY 365 DAVSPTWE--QLENAMVAVKTVVHGLVDFIQNYSRKGHETPOQPNKYKTSKMRDLRQQ 422
DB 246 QSVGGTIEAKALDKLTSIKIFEMLSIF-----DSKIK----- 279
QY 423 GGCPRGTNCTFAHSEBLEKYLRLNKKINATVTRTPLLNKGVNNTVTYTAGNVISVIGS 482
DB 280 -----ENSTMDLMBEVKK-----AEKISKVY----- 301
QY 483 TETTKVIPSTNGISNAENS-----VSQISRSTDSLRALETVKVKVGANGQNA 535
DB 302 SAGGRVDVMLTAISAFNSYNDPEKLEQMGFQID-----EPTGSVNQI----- 347
QY 536 GPSADSVTENKIGSPKTPVSNVAATSAGPSNVGTSLNSVPQKSSPFLTRVPVYPHSEN 595
DB 348 -PASEGGAER-----VNEVAAPAAG-----RPIPNARNLRGALNHAQVARN 389
QY 596 IQYFQDPRTOIPFEVQVPTQGYPPPPVTPVAGVAPCFVRFRVRSNNVPSSSLPPASMPYA 655
DB 390 -----RNRQARGVQQQDLPLPLPPPP-----HQLVMPPPAMPMP 424
QY 656 DHYST-----FSR-DMNASSPQP-----PPQVGPVPPVPSPGYAPV 694
DB 425 QHMQAORLKGFPQGRMN-----QPMMAQVGRPGQVIMGHPPPMHMQOOGPRMMY--M 478
QY 695 YDSRRIRWRPMPYQORDIIRSNLSLPMDMVHSSVYQTSLRERNYSLDGYYSVACQPPSEPR 754
DB 479 HQQHH-----PMHQVED-----QMQRQHRMQQPNMN--NIMHGNMQIAFQOHMNN 523
QY 755 TTVPILPREPCGHLKTSCEQIIRKPDQWAOY-----HTQKAPLVSSTLPA 800
DB 524 ANVPMPPQPAHPQ---QQQQQRHFNQGGFOGPPPPPPQRIHQHQOQOQYHQQLQVQ 580
QY 801 TQSPFP-----SPLFSVDFA-----DF-----SEVSQTKFEEDHLSHYSP-- 838
DB 581 QQQQFPQAPRFPPTQQQQQLQQLQAMNDQRFQQLADEIVAEQMVQNAVDAAPAD 640
QY 839 -----W-----SCGTIG-----SCINAIDSEPKOV-----IANSNAVIMDLSDGVKR 876
DB 641 IPRADWEEIQVAQRGAGDGFINFHPPIVEEVEEQPEVDQVYVNGQVVENVRAGDINN 700
```

```
QY 877 RVHLFETQRTKEDPII--PFSQDPIISKWGAISRSSRTGVHTTDPVQATASQGSATKP 934
DB 701 VIHF-----EEVPVLEPFDVGVNDHNGIENNVVEEQEQPDD-----VP 740
QY 935 ISVSDYVPYVNAVDSRWSSYGNEATSSAHYVERDRFIVTDLSGHRKHSSTGDLLSLELQQ 994
DB 741 PQINNFEV-----DDMREQVDAIEDEEF-----EPVLHR-----AAIEFOE 777
QY 995 AKSNLSLLQ--REANALAMQOKWNSLDEG--RHILTLLLSKEIELRNGELQSDYTEDATD 1050
DB 778 AADQDALBQDYRERNLLDHEAENEEDDDGLPYBEEDDEEEVHVRNSVVR--RVQRPRD 836
QY 1051 TKPDRDIELELSAL-----DDEPDGQSEPIEILDIOLGISSONDQLNGHVAENG 1102
DB 837 AAHISEVTRTSARLARNSNNTDEPSTSS-----GAAGPSN 874
QY 1103 HPVQOQKQEPKQKKQ 1118
DB 875 RP--PHSPTPPGNGKR 888

RESULT 10
C29356
hydroxyproline-rich glycoprotein (clone Hyp2.13) - kidney bean (fragment)
C:Species: Phaseolus vulgaris (kidney bean)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C:Accession: C29356
Mol. Cell. Biol. 7, 4337-4344, 1987
R:Corbin, D.R.; Sauer, N.; Lamb, C.J.
A:Title: Differential regulation of a hydroxyproline-rich glycoprotein gene family in wo
A:Reference number: A29356; MUID:88142825; PMID:3437892
A:Accession: C29356
A:Molecule type: mRNA
A:Residues: 1-368 <COR>
A:Cross-references: UNIPROT:Q09085; EMBL:M18095; NID:g169348; PIDN:AAA33765.1; PID:g1693
A:Experimental source: cv. Kievitsboon Koekoek
C:Superfamily: hydroxyproline-rich glycoprotein
C:Keywords: glycoprotein

Query Match      3.1%; Score 191; DB 2; Length 368;
Best Local Similarity 24.3%; Pred. No. 0.0028;
Matches 93; Conservative 30; Mismatches 153; Indels 106; Gaps 16;

QY 468 TVTTT--AGNVISVIGSTETTKLVPTNGI-----SNAENSVSQL 506
DB 3 TVEVTEVGGTKISYGNTKNGKYSITVEGLDYYKVGTVCKAQLHAPPKSGSRCSIPTK 62
QY 507 ISRSSTDSTLRALETVKVKVGKVGANGQNAAGPSADSVTENKIGSPKTPVS--NVAATSAGP 565
DB 63 LNEGTKLALKSKDYEVVLK-----AKPFAYA-----PKPYDCEKHKHSPTP 105
QY 566 SNVGTSLNSVP---QKSSPFLTRVPVYPHSENIQYFQDPRTOIPFEVQVPTQGYVPP 621
DB 106 YHKPYVYNSPPPPYVYKSPPPYKSPPPSPSPYKSPPP--PHKDPYPPYVYKSP 163
QY 622 PPTVAGVAPCVPRVRSNNVPSSSLPPASMPYADHYSTFSPDRDMNSSPY---QPPPPQ 678
DB 164 PPPSP---SPPPYVYKSPPPSPSPPP---PY--YKSPPPPPDPPSPPPYVYKSPPPPS 215
QY 679 PYGVPV---PVPSGMVAPVYDSRRIRWRPMPYQORDIIRSNLSLPMDMVHSSVYQTSLR 733
DB 216 ESPPPYVYKSPPPSPSPSP---PPYVYKSPPPSPSPSPSPSPSPPPYVYKSP--- 264
QY 734 ERYNSLDGYYSVACQPPSEPRTTVPL-----PREPCGHLKTSCEQIIRKPDQWAOYHT 787
DB 265 -----PPPDGPPPPYVYKSPPPSP-----SPPPPYVYKSP 296
QY 788 QKAPLVSSTLPAVATQSPTPPSP 809
DB 297 PPPSPSPPPPPYVYKSPPPSP 318
```

RESULT 11  
T20241  
hypothetical protein H05L14.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T20241; T23049  
R:Wilkinson, J.  
submitted to the EMBL Data Library, June 1996  
A:Reference number: Z19241  
A:Accession: T20241  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-2160 <WIL>  
A:Cross-references: UNIPROT:O17709; EMBL:Z75533; PIDN:CAA99823.1; GSPDB:GN00019; CESP:H05L14.2  
A:Experimental source: clone C5404  
R:Barlow, K.  
submitted to the EMBL Data Library, October 1997  
A:Reference number: Z19662  
A:Accession: T23049  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-2160 <W12>  
A:Cross-references: EMBL:Z99772; PIDN:CAB16922.1; GSPDB:GN00019; CESP:H05L14.2  
A:Experimental source: clone H05L14  
C:Genetics:  
A:Map position: 1  
A:Introns: 49/3; 130/3; 190/3; 237/3; 290/1; 463/3; 507/1; 524/2; 545/1; 636/2; 663/2; 7

Query Match 3.1%; Score 191; DB 2; Length 2160;  
Best Local Similarity 18.9%; Pred. No. 0.041;  
Matches 233; Conservative 162; Mismatches 447; Indels 394; Gaps 59;

QY 45 LHRKACPFQTAINTDIDL-----PVNFALLQLVCAQ-----VPDQHSIK 85  
DB 494 VHFQICP-----PTNEDLHLISRLKLSNLKRYSGNGFTSSGAERKGNIPDKLLDS 549

QY 86 LSNLIG-----ENKHY-----EVAKCKVEDLALYKPLSGGKGVASLNQSAISRPMORKLVTL-- 137  
DB 550 PSNMGSDSESDNPFSENEQC-----SVSSADKGISSNSQT--SDVPKGVITLDP 599

QY 138 -----VNCOLVEEGRVAMRAARSLGRTVTTELILQHONPOOLGANLWAAVRARGCOF 191  
DB 600 GTVCVSLTWKISEE-----SHKLG-----ELEAKKENDQI----- 631

QY 192 LGPAMQBEALKVLLALEDGSAISRKVLVLFVQRLEP-----RFPQASKTSIGHVV 243  
DB 632 -----LEPSTSVLFTERTAPPVVDNRSPTRSPRFPKSKSN----- 664

QY 244 QLLYRASCPIKTRDEBDSLMLQKEFRSVEALRREHDAQIVHIAEAGLRISPEQWSSL 303  
DB 665 -LITRLNVNEI--HSLKSLMDLLKQEKMSQEQRDDSDVD-----DESSVESSESPTSS 717

QY 304 LYGDLAHSHMQSIIDKLOSPESFAKSVQELTIVLQR--TGDPA-----NLNLRPHLEL 356  
DB 718 ITTISVASPTDPEAKRQKKEKKARQLSKQKQRLERAEAAKQRELNLAEHKKHLAK 777

QY 357 LANIDPNDAV-----SPTWEQLENAMVAVKTVVHGLVDFTQNYSRKGHETP 403  
DB 778 YAALVAKPTVTKTDEQFNKKSTQLQPRVEKTEKGNISVSVSSDSQKVE-----NRTG----- 828

QY 404 QPQNSKYKTSMCRDLRQCGGCRPTNC--TPAHSQEELEKYRLRNKKNATVTRTPPLN 461  
DB 829 -----SFGSPSTRLSNSNIP-----SCFATPSPDPSQLQSFVNSNDLDA----- 870

QY 462 KVGNNVTVTAGNVISVIGSTETTKIVPSTNGISNAENSVSQLISRSTDLRALETV 521  
DB 871 DTSVNLISITTDGVHLMVCVDKA-AQGVVVEATH-VGESEN-----TIEVVEDA 916

QY 522 KVKGVKGVANGAAGSADSVTENKIGSPPKTPVSNVVAATSAG--PSNVGTENLSPQKS 579  
DB 917 EKSFEILVCCQRR-----KVT-----PFTPRRKQSSSGDEQPASESTSAFDIPEHP 963

QY 580 S---PFLTRVPVYPHSENIQYFQDPRTQIPPEVQY----PQTGYPPPPPTYPAGVAPCV 633  
DB 964 PVIVFPRNAPIVYSSQTSNI-----PQIHPLPFRFLAMPQFSGHPPPPVIMG----- 1012

QY 634 PRFVSNVNPESLPPASMPYADHYSTFSPRDMNSSPYQPPPPOPYGVVP--PV----- 686  
DB 1013 PPIIIRPFQAP--PPQFIPTQSNYG-----RPPQFPPIGVVPIPIQLPQW 1055

QY 687 ----PSGMVA--PVYDSRRIRPPMYQ-----RDD-----ITRSN-----SLP 718  
DB 1056 RMSGPPVMMRGPPVYCSQPVPMHPMQNVQNMSSSRQDFGNVRLIRPNQKPLEITAP 1115

QY 719 PMDVHSSVYQ-----TSRERYNSLDGYSVACQPSSEPRTV 757  
DB 1116 PVSASNSKIEKMGKVRPEPRKEIINVSRVQVRNNGKOLDQF-----KTQKOTV 1167

QY 758 PLPREPCGHLKTSCEQIRRKPDQWQYHTQKAPLVSTLPVATQSPPTPPSLFSDVDFRA 817  
DB 1168 PETRKPINECKESEKDI-----VAKDGVVEQSEIVSQPENSXDL-- 1209

QY 818 DFSESVSGTKFEEDHLSHYSPWSCGTIGSCINAIDSEPKDVIANSNAVIMLDLSDGVKRR 877  
DB 1210 ----STSANKETSIVEQVAE-----KSLIAEKPDVRRSSS---IDL-NGD---- 1247

QY 878 VHLFETQRTKEEDPIIPFSDGPIISKWGAISRSTGHTTDPVQATASQCSATKPISV 937  
DB 1248 --LIEAVATVLDD---SDDESPIPS-----TSVSSSVLDSQTQSEAQDAINRVLGF 1295

QY 938 SDYVYVNAVDSRWSSYNGEATSSAHYVERDFIVTDLSGHRKHSSTGDLSSLLELQOAKS 997  
DB 1296 SKQCBETNF--SIWTFGENECPPEVIDI--NDNVLAAILPDE-----EPTTEE 1341

QY 998 NSLLQLQREANALAMQKWSLDEGRHLTLNLLSKEIELANGELQSDYTEDATDTKPDRI 1057  
DB 1342 NDVYTE-----IEDKARDFCFQKPTLLTTCQVEAYFSEHHDYVGSVD-----DM 1389

QY 1058 ELELSALDT---DEPDGQSEPIEELDQLGISSNDQLNG-----MAV 1099  
DB 1390 ELRKIFKETYGGISIPEDHKQIIEEVCSFQPGVIHVFDALIGSKLEFTYTIMSNIDRKIEK 1449

QY 1100 ENGHVPQHQKPPKQKQSLGEDHVLLEEQTILP 1135  
DB 1450 ENDRLYQHYSSRPADV-----GVENFYERFEKEILP 1480

RESULT 12  
F86387  
probable Pto kinase interactor [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C:Accession: F86387  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso  
ansen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: F86387  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-760 <STO>  
A:Cross-references: UNIPROT:O9C660; GB:AE005172; NID:gl1079512; PIDN:AG29223.1; GSPDB:  
C:Genetics:  
A:Map position: 1

Query Match 3.1%; Score 190.5; DB 2; Length 760;  
Best Local Similarity 25.4%; Pred. No. 0.0089;  
Matches 85; Conservative 29; Mismatches 116; Indels 105; Gaps 15;





**This Page Blank (uspto)**

APPLICANT: Hilbun, Erin  
APPLICANT: Donoho, Gregory  
APPLICANT: Turner, C. Alexander Jr.  
TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides  
FILE REFERENCE: Encoding the Same  
CURRENT FILING DATE: 2001-05-14  
PRIORITY FILING DATE: 2000-05-19  
NUMBER OF SEQ ID NOS: 64  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 48  
LENGTH: 1939  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(1939)  
OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-854-856-48

Query Match 3.1%; Score 193.5; DB 4; Length 1939;  
Best Local Similarity 19.2%; Pred. No. 7.3e-06;  
Matches 256; Conservative 168; Mismatches 527; Indels 379; Gaps 57;

QY 114 GGGVASTNQSALSRRMOKLVLYNLCOLVBEGRVR-AMRAARSILGERTVTTELILQHON 172  
DB 660 GQFSSSLTGVSSSQBIHQHPOQOQGIQTAPQQTVOYLSOTSSEATTAQPSQPOA 719  
QY 173 POOLSANLWAAVRARCCQFLGA-----MQEALKLVLLALEDEGSALEKVLVLFVVQRL 227  
DB 720 PQVLPVQVAGKSTQGVQVAPAEVAVAPQATQPTTLASSVDSAHSD---VASGMSDG 776  
QY 228 EPRFPOAKSTSGHVQVLLYRASCVKYTKRDEBSSLMQLK-----EPRFSVEALRRHED 281  
DB 777 NENVPSSSGRHEGRITTKHYRKS-VRSRSRHEKTSRPKLILNVSNKGDVVVECOLETHN 835  
QY 282 AQIVHAMEAGLRISPEQWSSLLY-GDLAKHSHQSIIDKLOSPESFAKSVQELTIVLQR 340  
DB 836 RKWTFKFDLD-GDNPEEITATIMVNDFFILATERESFVDQVREITEKADEMLSEDSVSEP 894  
QY 341 TGDPAANLRLRPHLELLANID-----PNPDVAVPTWEQLENAMVAVKTV 384  
DB 895 EGDQG-----LESLOGKDDYGFSGSQKLEGEFKQIPASSMP--QIGIPTSLTQV 944  
QY 385 VHGLVDFIQYGRKHGHTPOQPN-----SKYKTSKCRDLRQGGCGPRGTNCTFAHS 436  
DB 945 VH-----SAGRRFIVSPVESRLRESKVPFSEITDTVAASATAQSPG-----MNLSHS 991  
QY 437 QBELEKVR-----LRNKKI-----NATVTRTP-----LLNKVGNVNTVTTA--- 473  
DB 992 ASLSLQQAQFSELRAQMTGEPNTAPNFSHTGTPFPVPPFLSSTAGVPTTAAATAPVP 1051  
QY 474 -----GNVISVIGSTETTKGIVPNTGKISNAENS-----VSQLS 508  
DB 1052 ATSSPPNDISTSVIQEVT---VPTTEGIAGVATSTGVVTSGLPIPPVSEFVLSVVS 1108  
QY 509 RSTDSLRLALETVKYGVKYGAN-----QONAAGPSAD-SVTENKIG-----SPPKTPVSN 557  
DB 1109 SITIPAVVISITSPSLQVPTSTSEIWSSTALYPSVTVSATASAGGSTATPGKPPAV 1168  
QY 558 VAATSAGPNVGTLSNV-----PQKSSPLTRVVPVPHSENIQVF 599  
DB 1169 VSQQAAGSTVGTATLSVSTTSPFSTASQLSLSSSTSTPTLAETVVVVAHSLED-KTS 1227  
QY 600 QDPRTQIPREVQYPTGYPPPP-----TVPAGVAPCV-PRFVRNNVPESLPPASMP 653  
DB 1228 HASTTGLAFL-SAPSSSSPGAGVSSYISQPGGLHPLVIPSVIASITPILPQAGTSTP 1286  
QY 654 YADHYTSPRDMNSSPQPPPPQYGPVPPVPSGM-----YAPVYDS 697  
DB 1287 L-----LPQVPSIPLVQPVANVPAVQQTILHSQPOPALLPNQPHTHCPEVDS 1334

QY 698 RRIWRPMYQRDDI-----IRS--NSLPPMDVMHSSV-YQTSRERYNSLDGYYSVA-- 746  
DB 1335 DT--QPKAGIDDIKTLEKRLSLFSEHSSSSGAQHASVLETSLVIESVTVPGIPPTAVA 1392  
QY 747 -----CQPPSE-PRTTVPPLPREPC---GHLKTSCEBQIRR-KPDQWAQVHTQCA 790  
DB 1393 PSKLLTSTTCLPPTNLPLGTVALPVTVPVPGQVSTPVSTTSGVKFGT-----APSKP 1448  
QY 791 PLVSS-TLPVATQSPPTPSPPLFSVDFRADFSVSQTKFEEDHLSHYSPWSCGTIGSCIN 849  
DB 1449 PLTKAPVLPVGTGELP-----AGTLPSS--E 1470  
QY 850 AIDSEPKDVIANSNAVLMDLDS-----GDVKKRRVHLFETQRTK----- 888  
DB 1471 QLPPFPGLTQSQOQPLEDLDAQLRRTLSPEXITVTSVAVGPVMAAPTAITEAGTQPKG 1530  
QY 889 -----BEDPIIPSPDQPIISKWG-----AISRSSTGYHTTDPVQATASQGSATKPIVS 938  
DB 1531 VSQVKEGVLATSSGAGVFKMGRFQVSVAAQGAQKGNKSDAKSVHFESSTSESSVLS 1590  
QY 939 DIVPVYNAV-----DSRWSSYGNKETS-SAHYVVERDRFIVTDLS---G 977  
DB 1591 SSSPESTLVKPEPNGITIPGISSDVPESAHTKTASEAKSDTCQPTKVGRFQVTTANKVG 1650  
QY 978 HRKHSSTGDLSS-----LELQQAQNSLLQLQREANALAMQOKWN----- 1016  
DB 1651 RPSVSKTEDKITDTKKEGVPASPPFMDLEQAVLPAVIPKKEPELSEPSHLNGSPSDPEA 1710  
QY 1017 -----SLDEGR---HLTNLNSKELRN-----GELOSQY-TEDATDKPDRDIELEL 1061  
DB 1711 AFLSRDVEDGSGSPHQLSKSLPQNSLSQNSFNSSSYMSDNESDIEDEDLKLEL 1770  
QY 1062 SALDDEPDGQSEPIEILDIQLGISSQNDQLLN-----GMAVENGHPVQOHOKEPK 1114  
DB 1771 RRL-----RDHLKEIQDLSROKHEIESLYTKLGKVPVPAVIPPAAAPLSGRRRRPTK 1823  
QY 1115 QK-----KQSLGEDH-----VILEEQKTLIPVTSCFSPQLPVSISNASCLP 1155  
DB 1824 SKGSKSSRSSSLGNKSPQLSGNLSQSAASVLPQOQLHPGPN-----IPESQGNOLLOP 1878  
QY 1156 ITTSVSAGNL 1165  
DB 1879 LKPSPSDNL 1888

## RESULT 9

US-09-854-856-16  
Sequence 16, Application US/09854856  
Patent No. 6541252  
GENERAL INFORMATION:  
APPLICANT: Walke, D. Wade  
APPLICANT: Hilbun, Erin  
APPLICANT: Donoho, Gregory  
APPLICANT: Turner, C. Alexander Jr.  
TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides  
FILE REFERENCE: Encoding the Same  
CURRENT APPLICATION NUMBER: US/09/854,856  
CURRENT FILING DATE: 2001-05-14  
PRIORITY FILING DATE: 2000-05-19  
NUMBER OF SEQ ID NOS: 64  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 16  
LENGTH: 1999  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(1999)  
OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-854-856-16





341 TGDPAANLRLRPHLELLANID-----PNPDVSPTWQLENAMVAVKTV 384  
895 EGDQG-----LESLOKDDYGFSGSKLEGEFKQIPASSMP--QQIGIPTSSLTQV 944  
385 VHGLVDFFIONYSKRGHETPOPOPN-----SKYKTSMDRLRQOQCGPRGTCTFAHS 436  
945 VH-----SAGRRFIVSPVESRLRESKVFSEITDVAASTAQSPG-----MNLSHS 991  
437 QBELEKYR-----LRNKKI-----NATVTRFP-----LLNKGVNNTVTFTA--- 473  
992 ASLSLQQAFAFSELRRAQMTGEGTAPPNFSHTGTPFPVPPFLSSAGVPTTAATAPVP 1051  
474 -----GNVISVIGSTETTKIVPSTNGISNAENS-----VSQLS 508  
1052 ATSSPPNDISTSVIOSEVT---VPTTEGIAGVATSTGVVTSGLPIPPVSESVLSVVS 1108  
509 RSTDSLRALETIVKVKGVKYGAN-----CQNAAGPSAD--SVTENKIG---SPKTPVSN 557  
1109 SITIPAVWSISTSPSLQVPTSTSEIVVSTALYPSTVTSATSASAGGSTATPGPKPPAV 1168  
558 VAATSAGPSNVGTELNSV-----POKSSPFLTRVPVPPHSENIQYF 599  
1169 VSQQAAGSTVGTATLSVSTSTTSPSTASQLSIQSSSTSTPLAETVVVSAHSLD-KTS 1227  
600 QDPRTQIPPEVPOYPOTGYPPPP-----TVPAGVAPCV-PRFVRNNVPESLPPASMP 653  
1228 HSTTGLAFLS-SAPSSSSPGAGVSSYISQPGHLPLVIPSIVASTPILPQAAGPTSTP 1286  
654 YADHYSTFSPRDMNSPYQPPPPQYGPVPPVPSGM-----YAPVDS 697  
1287 L-----LPQVSPILPLVQVANVPAVQOTLIHSQPQALLPNQPHTHCPEVDS 1334  
698 RIWRPMPYQRDI-----IRS--NSLPMDVMHSSV-YOTSLRERYNSLDGYSAV-- 746  
1335 DT--QKAPCIDDIKLEKLRLFSFSEHSSGAQHASLETSLVISTVTPGIPPTAVA 1392  
747 -----COPPE-PRTTVPLPREPC---GHLKTSCEQIRR-KPDQMAQVHTQKA 790  
1393 PSKLLTSTSTCLPPTNPLGTVLVPVTPVPGVSTPVSITTSVKPGT-----APSKP 1448  
791 PLVSS--TLPVATQSTPPSPPLSFVDFRADFSESVSGTKFEEDHLSHYSPWSCGTIGSCIN 849  
1449 PLTKAPVLPVGTLP-----AGTLPS--E 1470  
850 AIDSEPKDVIANNAVIMLDS-----GDYKRVHLFETQRTK-----888  
1471 QLPPFPGPSLTSQSQPLEDLDAQLRRLTSPXETIVTSAVGPVSMAPTATEAGTQPKG 1530  
889 ----BEDPIIPFSDGPIISKWG-----AISRRTGYHTTDPVQATASQSGATKPIVS 938  
1531 VSQVKEGVLATSSGAGVFKGRFQVSVADGAQKEGKSEDAKSVHFESETSESIVLS 1590  
939 DYVPVYNAV-----DSRWSSYGNBATS-SAHYVBRDRFIVTDLG---G 977  
1591 SSSPESTLVKPEPNEGITIPGISSDVPESAHTTASEAKSDTGQPTKVGRFOVTTTANKVG 1650  
978 HRKXSTGDLSS-----LELOAKNSLLLOREANALAMQKWN-----1016  
1651 RFSVSKTEDKITDITKKEGVPASPFPMDLEQAVLPAVIPKEKPELSEPHLNGFSSDPEA 1710  
1017 -----SLDEGR---HLTNLLSKIELRN-----GELQSDY--TEDATDKPRDIELEL 1061  
1711 AFLSRVDDGSGFPHQJLSSKLSFQNSLSQSLNSFNSNSNESDIEDIEDLKLLEL 1770  
1062 SALDTEBPQCGSEPIEILDIQLGISQNDQLLN-----GMAVENGHVPVOHQKEPPK 1114  
1771 RRL-----RDXHLKEIQDLSRQKHEIESLYTLKGVKVPVAVIIPPAAPLSGRRRRPTK 1823  
1115 QK-----KQSLGEDH-----VILEEQKTLPVTCFQOPLPVISNASCLP 1155  
1824 SKGSKSRSSSLGNKSPQLSGNLSGQSAASVLHPQOTLHPGN-----IPESQNLQLOP 1878  
1156 ITTSVSAGNL 1165

Db 1879 LKPSPSDNL 1888  
RESULT 11  
US-09-854-856-14  
; Sequence 14, Application US/09854856  
; Patent No. 6541252  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Hilbun, Erin  
; APPLICANT: Donoho, Gregory  
; APPLICANT: Turner, C. Alexander Jr.  
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides  
; TITLE OF INVENTION: Encoding the Same  
; FILE REFERENCE: LEX-0178-USA  
; CURRENT APPLICATION NUMBER: US/09/854,856  
; CURRENT FILING DATE: 2001-05-14  
; PRIOR APPLICATION NUMBER: US 60/206,015  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 2136  
; TYPE: PRN  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(2136)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-854-856-14  
Query Match 3.1%; Score 193.5; DB 4; Length 2136;  
Best Local Similarity 19.2%; Pred. No. 8.6e-06;  
Matches 256; Conservative 168; Mismatches 527; Indels 379; Gaps 57;  
QY 114 GKGVASLNQSAISPMQKLVTLVNCQLVEEGVR-AMRAARSIGERTVTTELILQHQN 172  
Db 720 GQPSSSSLTGVSSSQPIQHPOQQOQIQQTAPPOQTVOYLSQTSSTSEATTAQPVSPQA 779  
QY 173 POOLSANLWAVRARGCQFLGPA-----MOEALKLVLLALEDGSAISKVLVLFVQRL 227  
Db 780 POVLQVSGAKOSTQGVQVAPBVAQPOATQPTTLASSVDSAHSD---VASGMSDG 836  
QY 228 EPRFPOAKSTIGHVVQLLYRASCFTKTRDSDSLMOLK-----EERFSYEALRREHD 281  
Db 837 NENFSSSGRHGRTTKHYRKS-VRSRERHEKTRPKLRILNVNKGDRVVECOLETHN 895  
QY 282 AQIVHIANEAGLRISPEQWSSLLY-GDLAKHSHMOSIIDKLOSPESFAKSVQELTIVLQR 340  
Db 896 RKMVTFKFLLD-GDNPEEATIMVNNDFILATERESFVDQVREIIEKADEMLSEDSVSEP 954  
QY 341 TGDPAANLRLRPHLELLANID-----PNPDVSPTWQLENAMVAVKTV 384  
Db 955 EGDQG-----LESLOKDDYGFSGSKLEGEFKQIPASSMP--QQIGIPTSSLTQV 1004  
QY 385 VHGLVDFFIONYSKRGHETPOPOPN-----SKYKTSMDRLRQOQCGPRGTCTFAHS 436  
Db 1005 VH-----SAGRRFIVSPVESRLRESKVFSEITDVAASTAQSPG-----MNLSHS 1051  
QY 437 QBELEKYR-----LRNKKI-----NATVTRFP-----LLNKGVNNTVTFTA--- 473  
Db 1052 ASLSLQQAFAFSELRRAQMTGEGTAPPNFSHTGTPFPVPPFLSSAGVPTTAATAPVP 1111  
QY 474 -----GNVISVIGSTETTKIVPSTNGISNAENS-----VSQLS 508  
Db 1112 ATSSPPNDISTSVIOSEVT---VPTTEGIAGVATSTGVVTSGLPIPPVSESVLSVVS 1168  
QY 509 RSTDSLRALETIVKVKGVKYGAN-----CQNAAGPSAD--SVTENKIG---SPKTPVSN 557  
Db 1169 SITIPAVWSISTSPSLQVPTSTSEIVVSTALYPSTVTSATSASAGGSTATPGPKPPAV 1228  
QY 558 VAATSAGPSNVGTELNSV-----POKSSPFLTRVPVPPHSENIQYF 599

|    |      |              |              |           |            |           |              |        |      |          |        |     |       |      |        |     |     |       |        |        |        |      |   |   |   |   |   |     |   |   |   |   |        |        |      |     |   |   |   |        |       |      |   |      |   |   |   |     |   |      |   |      |      |      |
|----|------|--------------|--------------|-----------|------------|-----------|--------------|--------|------|----------|--------|-----|-------|------|--------|-----|-----|-------|--------|--------|--------|------|---|---|---|---|---|-----|---|---|---|---|--------|--------|------|-----|---|---|---|--------|-------|------|---|------|---|---|---|-----|---|------|---|------|------|------|
| Db | 1229 | VSQQAAGSTTGA | YLTLSVSTTTFP | SPSTASQLS | IQSSSTSTPT | LAETVVVSA | HSLD-KTS     | 1287   |      |          |        |     |       |      |        |     |     |       |        |        |        |      |   |   |   |   |   |     |   |   |   |   |        |        |      |     |   |   |   |        |       |      |   |      |   |   |   |     |   |      |   |      |      |      |
| Qy | 600  | QDRPTQIPFEV  | QVPQYQGYPP   | PPPP----- | TVPA       | GA        | VAPCV-PRFVRS | NNVP   | ESSL | PPASMP   | 653    |     |       |      |        |     |     |       |        |        |        |      |   |   |   |   |   |     |   |   |   |   |        |        |      |     |   |   |   |        |       |      |   |      |   |   |   |     |   |      |   |      |      |      |
| Db | 1288 | HSSTTG       | LA           | FSL-SAPSS | SSSPGAGV   | SSYISQ    | PGGLHL       | VL     | PS   | VI       | ASTPIL | POA | GTSTP | 1346 |        |     |     |       |        |        |        |      |   |   |   |   |   |     |   |   |   |   |        |        |      |     |   |   |   |        |       |      |   |      |   |   |   |     |   |      |   |      |      |      |
| Qy | 654  | YADHYST      | SPRDR        | MNS       | SPYPP      | PPQY      | GP           | PP     | VP   | SGM----- | YAP    | YDS | 697   |      |        |     |     |       |        |        |        |      |   |   |   |   |   |     |   |   |   |   |        |        |      |     |   |   |   |        |       |      |   |      |   |   |   |     |   |      |   |      |      |      |
| Db | 1347 | L-----       | LPQ          | ES        | PL         | PL        | Q            | P      | AN   | PA       | VQ     | TL  | HS    | Q    | P      | ALL | PNQ | PH    | TC     | PE     | VD     | 1394 |   |   |   |   |   |     |   |   |   |   |        |        |      |     |   |   |   |        |       |      |   |      |   |   |   |     |   |      |   |      |      |      |
| Qy | 698  | RRWR         | PMY          | OR        | DDI-----   | IRS--     | N            | S      | L    | P        | M      | D   | M     | V    | H      | S   | S   | V-YQ  | T      | S      | L      | R    | E | R | Y | N | S | L   | D | G | Y | S | V      | A--    | 746  |     |   |   |   |        |       |      |   |      |   |   |   |     |   |      |   |      |      |      |
| Db | 1395 | DT--Q        | AP           | G         | I          | D         | I            | K      | T    | L        | E      | E   | K     | L    | R      | S   | F   | S     | H      | S      | S      | G    | A | H | S | V | L | S   | T | S | T | V | T      | P      | G    | I   | T | T | A | 1452   |       |      |   |      |   |   |   |     |   |      |   |      |      |      |
| Qy | 747  | -----        | COP          | PS-PR     | T          | VL        | PRE          | C----- | G    | H        | L      | K   | T     | S    | C      | E   | E   | Q     | I      | R      | -K     | P    | D | W | A | Q | Y | H   | T | Q | K | A | 790    |        |      |     |   |   |   |        |       |      |   |      |   |   |   |     |   |      |   |      |      |      |
| Db | 1453 | PSK          | LLT          | T         | T          | T         | C            | L      | P    | T        | N      | L   | P     | L    | G      | T   | V   | A     | L      | P      | V      | P    | V | T | T | G | V | S   | T | P | V | T | T      | S      | T    | G   | V | K | P | T----- | APSKP | 1508 |   |      |   |   |   |     |   |      |   |      |      |      |
| Qy | 791  | PLV          | SS--         | T         | L          | P         | V            | A      | T    | O        | S      | P   | T     | P    | S      | P   | L   | F     | S      | V      | D      | R    | A | D | F | S | E | S   | V | G | T | K | F      | E      | E    | D   | H | L | S | H      | Y     | P    | S | M    | C | G | T | I   | A | S    | C | I    | N    | 849  |
| Db | 1509 | PL           | T            | K         | A          | P         | V            | L      | P    | V        | G      | T   | E     | L    | P----- | A   | G   | T     | L      | P      | S--E   | 1530 |   |   |   |   |   |     |   |   |   |   |        |        |      |     |   |   |   |        |       |      |   |      |   |   |   |     |   |      |   |      |      |      |
| Qy | 850  | AID          | SE           | P         | K          | D         | V            | I      | A    | N        | S      | A   | N     | A    | L      | M   | D   | L     | S----- | G      | D      | V    | K | R | V | H | L | F   | T | Q | R | T | K----- | 888    |      |     |   |   |   |        |       |      |   |      |   |   |   |     |   |      |   |      |      |      |
| Db | 1531 | QL           | P            | F         | P          | G         | S            | L      | T    | S        | O      | Q   | P     | L    | E      | D   | L   | A     | Q      | L      | R      | T    | L | S | P | E | X | I   | T | V | S | A | G      | V      | F    | S   | W | A | A | P      | T     | A    | I | T    | E | A | G | T   | O | P    | K | G    | 1590 |      |
| Qy | 889  | ----         | E            | E         | D          | P         | I            | P      | F    | S        | D      | G   | P     | I    | S      | K   | W   | ----- | A      | I      | S      | R    | S | R | T | G | Y | H   | T | T | D | P | V      | A      | T    | A   | S | O | G | S      | A     | T    | K | I    | S | V | S | 938 |   |      |   |      |      |      |
| Db | 1591 | VSQ          | K            | E         | G          | F         | V            | L      | T    | S        | G      | A   | G     | V    | F      | K   | R   | F     | Q      | V      | S      | A    | D | G | A | K | E | K   | N | K | S | E | D      | A      | K    | S   | V | H | F | E      | S     | T    | S | E    | S | S | V | L   | S | 1650 |   |      |      |      |
| Qy | 939  | DY           | P            | V         | N          | A         | V-----       | D      | S    | R        | W      | S   | S     | Y    | G      | N   | E   | A     | T      | S----- | S      | A    | H | V | V | E | R | D   | R | I | V | I | D      | L      | S--G | 977 |   |   |   |        |       |      |   |      |   |   |   |     |   |      |   |      |      |      |
| Db | 1651 | SS           | P            | E         | S          | T         | L            | V      | K    | P        | E      | N   | G     | I    | T      | P   | G   | I     | S      | S      | D      | V    | P | S | A | K | T | T   | A | S | E | A | K      | S      | T    | G   | O | P | T | K      | V     | R    | F | O    | V | T | T | A   | N | K    | V | 1710 |      |      |
| Qy | 978  | HR           | H            | S         | T          | G         | D            | L      | L    | S-----   | L      | E   | O     | A    | K      | S   | N   | L     | L      | O      | R      | E    | A | N | A | L | A | M   | A | Q | O | K | W      | N----- | 1016 |     |   |   |   |        |       |      |   |      |   |   |   |     |   |      |   |      |      |      |
| Db | 1711 | R            | F            | S         | V          | S         | K            | E      | D    | I        | T      | D   | T     | K    | E      | G   | P   | V     | A      | S      | P      | F    | M | D | E | Q | A | V   | L | P | A | V | I      | P      | K    | E   | K | E | P | S      | E     | L    | P | S    | H | L | G | P   | S | D    | P | E    | A    | 1770 |
| Qy | 1017 | ----         | S            | L         | D          | E         | G            | R--    | H    | L        | T      | N   | L     | L    | S      | K   | E   | I     | E      | L      | N----- | G    | E | L | O | S | D | Y-T | E | D | A | T | D      | T      | K    | P   | R | D | I | E      | L     | E    | L | 1061 |   |   |   |     |   |      |   |      |      |      |
| Db | 1771 |              |              |           |            |           |              |        |      |          |        |     |       |      |        |     |     |       |        |        |        |      |   |   |   |   |   |     |   |   |   |   |        |        |      |     |   |   |   |        |       |      |   |      |   |   |   |     |   |      |   |      |      |      |

RESULT 12

US-09-854-856-64

; Sequence 64, Application US/09854856

; Patent No. 654125;

**GENERAL INFORMATION:**

APPLICANT: Walke, D. Wade

**APPLICANT:** Hilbun, Erin

**APPLICANT: Donoho, Gregory**

APPLICANT: Turner, C. Alexander Jr.

; TITLE OF INVENTION: No. 65411

; TITLE OF INVENTION: Encoding the Same

; FILE REFERENCE: LEX-0178-USA

```

; CURRENT APPLICATION NUMBER: US
: CURRENT FILING DATE. 2001-05-

```

; CURRENT FILING DATE: 2001-05-14  
 ; PRIORITY APPLICATION NUMBER: US 60/206 016

;  
; PRIOR APPLICATION NUMBER: US 6  
; PRIOR FILING DATE: 2000-05-19

; PRIOR FILING DATE: 2000-0  
 : NIMBER OF SEQ ID NOS: 64

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 64

; LENGTH: 1911

```

;
; TYPE: PRT

```

ORGANISM: *Homo sapiens*

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1) .. (1911)

; OTHER INFORMATION: Xaa = Any Amino Acid

US-09-854-856-64

|                           |        |                    |             |              |
|---------------------------|--------|--------------------|-------------|--------------|
| Query Match               | 3.1%;  | Score 193;         | DB 4;       | Length 1911; |
| Best Local Similarity     | 19.3%; | Pred. No. 7.8e-06; |             |              |
| Matches 248; Conservative | 159;   | Mismatches 498;    | Indels 378; | Gaps 56;     |

Db 1436 -----AGTLPS---EQLPPFPGLSQTOSQOQPLEDLDAQLRRLTSPEXITVTSAVGPVMSMAAP 1489  
QY 880 LPEQRTTK-----EDPIIPFSDGPIISKWG-----AISRSRTGYHTTDPVQAT 925  
Db 1490 TAITAGTQPKGVQKGVKGVPLATSSGAGVFMGRFQVSVAAADGAQKGNKSEDAKSV 1549  
QY 926 ASQGSATKPIVSVDYVYVNAV-----DSRWSSYGNBATS-SAHYVER 967  
Db 1550 HFESSSTSESSVLSSSPSTLVKPEPNGITIPGISSDVPESAHTKTASEAKSDTQPTKV 1609  
QY 968 DRFIYTDLS---GHRKHSSTGDLSS-----LELOQAKNSLLQREANALAM 1011  
Db 1610 GRFQVTTANKVGRFVSFKTEKIDTDTKKEGVPASPPFMDLQAVLPAVIPKEKEPSE 1669  
QY 1012 QQKWN-----SLDEGR---HLTLNLSKEIELRN-----GELOSDY-TEDA 1048  
Db 1670 PSHLNGPSSDPAAFLSRDVGSGSPHPSQKLSQSLNSQSLNSFNSSYMSDN 1729  
QY 1049 TDTKPDRIEELSALDTPDQSQSEPIEELDIQLGISQNDQLLN-----GMAVEN 1101  
Db 1730 ESDIEDDLKLELRL-----RDHKLKEIQDLSQKHEIESLYTLKGVPPAVIIPP 1782  
QY 1102 GHPVQHQKEPKQK-----KQSLGEDH-----VILEEQKTLPTVSCFSQ 1142  
Db 1783 AAPLSGRRRPTKSGKSSRSLSGNKSPQLSGNLSQSAASVLPQOQLHPGPN-----1838  
QY 1143 PLPVSISSNACLPITTSVAGNL 1165  
Db 1839 -IPESGQNLQPLKPSDNDL 1860

## RESULT 13

US-09-854-856-32

; Sequence 32, Application US/09854856

; Patent No. 6541252

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Hilbun, Erin

; APPLICANT: Donoho, Gregory

; APPLICANT: Turner, C. Alexander Jr.

; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides

; FILE REFERENCE: Encoding the Same

; CURRENT APPLICATION NUMBER: US/09/854,856

; CURRENT FILING DATE: 2001-05-14

; PRIOR APPLICATION NUMBER: US 60/206,015

; PRIOR FILING DATE: 2000-05-19

; NUMBER OF SEQ ID NOS: 64

; SOFTWARE: Fast-Seq for Windows Version 4.0

; SEQ ID NO 32

; LENGTH: 1971

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)---(1971)

; OTHER INFORMATION: Xaa = Any Amino Acid

US-09-854-856-32

## Query Match

Best Local Similarity 3.1%; Score 193; DB 4; Length 1971;

Matches 248; Conservative 159; Mismatches 498; Indels 378; Gaps 56;

QY 160 ERTVTEILQHPQQLSANLWNAVRARGCQFLGPA-----MQEBAKLVLLALEDGSAL 214

Db 739 EATTAQVSPQAPQVLPQVSAQKQSTQGVSAPEAVAPQATQPTTLASSVDSAH 798

QY 215 SRKVLVTVQRLPFPQASKTSIGHVQLLYRASCVKYTKEDDESLMLQK-----E 268

Db 799 SD---VASGMSDGNENPVSSGHEGRTTKRHYRKS-VRSRSRHEKTSRPLKILNVSNK 854

QY 269 EPRSVALREHDAQIVHIAEAGLRISPEWSLLY-GDLAKSHMQSIIDKLQSPESF 327

Db 855 GDRVVECOLETHNRKMTFKFDLD-GDNPEIATIMVNDPILATERESFVDQVREIEIK 913  
QY 328 AKSVOELTVLQRTGD PANLRLRPHLELLANID-----PNPDVSPMT 371  
Db 914 ADEMLSDVSVEPEGDQ-----LESLOKDDYGFSGSKLEGEFKQPIPASSMP-- 963  
QY 372 EOLENAMAVKTVHGLVDFIQNYSRKHEHETPOQPN-----SKYKTSMDRLRQOG 423  
Db 964 QOIGIPTSLTQVHV-----SAGRRFIVSPVPSRLRESKVFSEITDVTAASTAOSP 1016  
QY 424 GCPRTGCTFAHSQSELEKYR---LRNKKI-----NATVTRTPP-----LLNKV 463  
Db 1017 G-----MNLSSASSLSLQOAFSELRAQMEGTAPPHSHTGTFPVVPPFLSSIA 1070  
QY 464 GVNNTVTITTA-----GNVISVIGSTETTKIPIVSTNGISNAENS-----502  
Db 1071 GVPTTAATAVPATSPNDISTSVIQSEVT---VPTTEGIAGVATSTGVVTSGLPIP 1127  
QY 503 -----VSQILSRSTDTLRALETIVKVKGVKAN-----GQNAAGPSAD-SVTENKIG 548  
Db 1128 PVSESPVLSSVVSSITIPAVVSIISTSPSLQVPTSTSEIVVSSSTALYPSVTVSATASAG 1187  
QY 549 ---SPPKTPVSNVAATSAGPSNVCTELNSV-----POKSSSPLTRV 586  
Db 1188 GSTATPGKPPAVVQQAGSTTVGATLTSVSTTSPSTASQSLQSLSSSTSTPTLAET 1247  
QY 587 PVPYPPHSENIQYFQDPTQIPFEVPOYQTYGYPPPP-----TVPAGVAPCV-PRFVRN 640  
Db 1248 VVVAHSLD-KTSHSSTTGLAFSL-SAPSSSSSPGAGVSSYISQGLHPLVIPSVIAT 1305  
QY 641 NVPESSLPASMPYADHYSTSPRDRMNSSPVQPPPPQYGPVPPVPSGM-----690  
Db 1306 PILQQAAGTSTPL-----LPQVPSIPPLVQFVANVPAVQOQLIHSQOPALL 1353  
QY 691 -----YAPVDSRRIMWPPMYQRDDI-----IRS--NSLPPMDVMHSSV-YOTSLRER 735  
Db 1354 PNQPHTHCEVDSDT--QPKAPGIDDIITLBEKLKSLFSEHSSSQAQHASVLESLSLVE 1411  
QY 736 YNSLDGYTSA-----CQPPSE-PRTTVPLPREPC---GHLKTSCEEQIRR- 777  
Db 1412 SITVTGITTAVAPSKLLTSTTCLPPLNPLGTVLPAVPTVPTVPGQVSTVSTTSGV 1471  
QY 778 KPDQWAQYHTQKAPLVSS-TLPVATQSPPTPSPPLSFVDFRADFSESVCSTKEEDHLSHY 836  
Db 1472 KPGT---APSKPPLTKAPVLPVGTPLP-----1495  
QY 837 SPWSCGTIGSCINADSEPKDVIANNAVLMDLDS-----GDKKRRVH 879  
Db 1496 ----AGTLPS---EQLPPFPGLSQTOSQOQPLEDLDAQLRRLTSPEXITVTSAVGPVMSMAAP 1549  
QY 880 LPEQRTTK-----EDPIIPFSDGPIISKWG-----AISRSRTGYHTTDPVQAT 925  
Db 1550 TAITAGTQPKGVQKGVKGVPLATSSGAGVFMGRFQVSVAAADGAQKGNKSEDAKSV 1609  
QY 926 ASQGSATKPIVSVDYVYVNAV-----DSRWSSYGNBATS-SAHYVER 967  
Db 1610 HFESSSTSESSVLSSSPSTLVKPEPNGITIPGISSDVPESAHTKTASEAKSDTQPTKV 1669  
QY 968 DRFIYTDLS---GHRKHSSTGDLSS-----LELOQAKNSLLQREANALAM 1011  
Db 1670 GRFQVTTANKVGRFVSFKTEKIDTDTKKEGVPASPPFMDLQAVLPAVIPKEKEPSE 1729  
QY 1012 QQKWN-----SLDEGR---HLTLNLSKEIELRN-----GELOSDY-TEDA 1048  
Db 1730 PSHLNGPSSDPAAFLSRDVGSGSPHPSQKLSQSLNSQSLNSFNSSYMSDN 1789  
QY 1049 TDTKPDRIEELSALDTPDQSQSEPIEELDIQLGISQNDQLLN-----GMAVEN 1101  
Db 1790 ESDIEDDLKLELRL-----RDHKLKEIQDLSQKHEIESLYTLKGVPPAVIIPP 1842  
QY 1102 GHPVQHQKEPKQK-----KQSLGEDH-----VILEEQKTLPTVSCFSQ 1142  
Db 1843 AAPLSGRRRPTKSGKSSRSLSGNKSPQLSGNLSQSAASVLPQOQLHPGPN-----1898

QY 1143 PLPVSIASCLPITTSVSAGNL 1165  
Db 1899 -IPESQNLQLOPLKPSFSDNL 1920

RESULT 14  
US-09-854-856-62  
; Sequence 62, Application US/09854856  
; Patent No. 6541252  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Hilbun, Erin  
; APPLICANT: Donoho, Gregory  
; APPLICANT: Turner, C. Alexander Jr.  
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides  
; FILE REFERENCE: Encoding the Same  
; CURRENT APPLICATION NUMBER: US/09/854,856  
; PRIOR FILING DATE: 2001-05-14  
; PRIOR APPLICATION NUMBER: US 60/206,015  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 62  
; LENGTH: 2048  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(2048)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-854-856-62

Query Match 3.1%; Score 193; DB 4; Length 2048;  
Best Local Similarity 19.3%; Pred. No. 8.8e-06;  
Matches 248; Conservative 159; Mismatches 498; Indels 378; Gaps 56;

QY 160 ERTVELILQHNPQOLSANLWAVRACQFLGA-----MQEALKLVLLALEDGSA 214  
Db 679 EATTAQVSPQAPQVLPQVSAQKOSTQGVSAFAEPVAVAPQATQPTTLASSVDSAH 738  
QY 215 SRKVLVLFVQRLPFRPOAKSTSGHVVLRYRACFKVTKRDESDSLMLK-----E 268  
Db 739 SD---VASGMSGDNVPSSSGRHGRTRKYRKS-VRSRSHKTSRPKRLILNVNK 794  
QY 269 EFRSVEALRREDAQIVHTIAMBAGLRISPEQWSSLLY-GDLAKHSMOSIIDKLOSPESP 327  
Db 795 GDRVVECOLETHNRKMTVTFKFLD-GDNPEETATIMVNDFFILATERESFVDQVREIEK 853  
QY 328 AKSVQELTIVLQRTGDPANLRLRPHLELLANID-----PNPDVSPTW 371  
Db 854 ADEMLSEDVSPPEGDQ-----LESLOGKDDYGFSGSKLEGEFKQPTIPASSMP-- 903  
QY 372 EOLENANWAKTVTHGLVDFIYGRKKGHETPOPOPN-----SKYKTSRCRLROQG 423  
Db 904 QQIGTPTSLTQVH-----SAGRRFIVSPVPSRLRESKVFSEITDVAASTAASP 956  
QY 424 GCPRGNTCTFAHQBELEKYR-----LRNKI-----NAVRTFP-----LLNKV 463  
Db 957 G-----MNLSSHASLSLQAFSELRRRAQMTGPNAPPNFSHTGPTFPVVPPLSSIA 1010  
QY 464 GUNNVITTTA-----GNVISIGSTETTKGVLPSTNGISNAENS----- 502  
Db 1011 GVPTTAAATAPVATSPSPNDISTSVIQSEVT---VPTTEGIAGVATSTGVVTSGLPPIP 1067  
QY 503 -----VSOLISRSRSTSLRALETVKVKGVKAN-----QONAAGPSAD-SVTENKIG 548  
Db 1068 PVSESPVLSVSSITIPAVVISITSPSLQVPTSTSEIVWSSTALVPSVTTSATASAG 1127  
QY 549 -----GPKPTPVSNVAATAGSPNVGTENSV-----PQKSSPPLTRV 586  
Db 1128 GSTATPGPKPPAVVSQQAAGSTTVGATLTLSVSTTTSFPTSTASQLSITLSSTSTPTLAET 1187

QY 587 PVYPHSENIQYFQDPRTOIPEVPPQYPTQGYVPPPP-----TVPAGVAPCV-PRFVRSN 640  
Db 1188 VVSAHSLD-KTSHSSTTGLAFSL-SAPSSSSPCAGVSSYISQPGGLHPLVIPSVIAS 1245  
QY 641 NVPESSLPASMPYADHYSTFSPRDRMNSPYQPPQPPQYGPVPPVPSGM----- 690  
Db 1246 PILQOAGPTSTPL-----LPQVPSIPLVQPVANVPAVQOTLIHQSPQAPALL 1293  
QY 691 -----YAPVYDSRRIWRPPMYQRDDI-----TRS--NSLPPMDVMHSSV-YQTSLE 735  
Db 1294 PNQPHTHCPVDSDT--QPKAPGIDDIKTLEKRLSLFSEHSSSSGAQHASVLESLVIE 1351  
QY 736 YNSLDGYYSA-----CQPPSE-PRTTVPLPREPC---GHLKTSCEBQIR- 777  
Db 1352 STVTPIGITTAVAPSKLTLTSTCLPPTNPLGTVALVPTVPTVPGQVSTPVTSTTSGV 1411  
QY 778 KPDQWAOYHTQKAPLVSS-TLPVATQSPTPSPPLFSVDFRADFSSESVTGTFEEDHLSHY 836  
Db 1412 KPGT---APSKPPLTKAPVLPVGTLP----- 1435  
QY 837 SPWSCGTIGSCINAIIDSEPKDVIANNAVLMDLDS-----GDVKKRVH 879  
Db 1436 ---AGTLPS--EQLPPFPFSLTQSQQPLEDLDAQLRLTSLPEXITVTSVAVGPVMAAP 1489  
QY 880 LFEQRTK-----EEDPIIPFSDGPIISKWG-----AISRSRTGYHTTDPVOAT 925  
Db 1490 TAITAQTQOKGVSOVKEGPVILATSSGAGVFKGRFOVSVADGAQKEGKNKSEDAKSV 1549  
QY 926 ASQGSATKPIVSVDYVYVNAV-----DSRWSSYGNPEATS-SAHYVER 967  
Db 1550 HFESSTSESSVLSSSPESTLVKPEPNGITIPGISSDVPESAHKTTASEAKSDTQPTKV 1609  
QY 968 DRPIVTDLS---GHRKHSSTGDLLS-----LELQAKSNLLOREANALAM 1011  
Db 1610 GRFOVTTANKVGRFSVSKTKITDTKKEGVASPPFMDLEQAVLPVAPKKEPELSE 1669  
QY 1012 QOKWN-----SLDEGR---HLTNLLSKELELN-----GELQSDY-TEDA 1048  
Db 1670 PSHLNGSPSDEAPFLSRDVGSGSPHSPHQLSKSLPSONLSQSNSSNFSNWSN 1729  
QY 1049 TDTKPDRIELELSALDTDEPDGQSEPIEIEDIQLGISSQNDQLN-----GMAYEN 1101  
Db 1730 ESDIEDLKLRL-----RDKHLKEIQLSRQKHEIESLYTKLGKVPVAVIIPP 1782  
QY 1102 GHPVOHQKPEPKQK-----KQSLGEDH-----VILEEQKTLPTVTSQFSQ 1142  
Db 1783 AAPLSGRRRRTKSGKSSRSSSLGNKSPQLSGNLSCQSAASVLHPQOTLHPGPN---- 1838  
QY 1143 PLPVSIASCLPITTSVSAGNL 1165  
Db 1839 -IPESQNLQLOPLKPSFSDNL 1860

RESULT 15  
US-09-854-856-30  
; Sequence 30, Application US/09854856  
; Patent No. 6541252  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Hilbun, Erin  
; APPLICANT: Donoho, Gregory  
; APPLICANT: Turner, C. Alexander Jr.  
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides  
; FILE REFERENCE: Encoding the Same  
; CURRENT APPLICATION NUMBER: US/09/854,856  
; PRIOR FILING DATE: 2001-05-14  
; PRIOR APPLICATION NUMBER: US 60/206,015  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 30

Query Match 3.1%; Score 193; DB 4; Length 2108;  
Best Local Similarity 19.3%; Pred. No. 9.2e-06;  
Matches 248; Conservative 159; Mismatches 498; Indels 378; Gaps 56;  
160 ERTVTELIHQHNPQOLSANLWAAVRACQCLGPA-----NQEALKVLVLALEDGSA 214  
739 EATTAQVSPQAPQVLPQVSAGKQSTQGVAPAEPAVAPQATQPTTLASSVDSAH 798  
215 SRKVLVFLVQLEPRFPQAKTSIGHVVQLLYRASCFTVKRDEDSLMQJL-----E 268  
799 SD---VASGMDGNEVPSSGRHEGRTTKRHYRKS-VRSRSHKTSRPLRLNVSNK 854  
269 EFRSYEARREDAQIVHAMEAGLRISPEQWSSLLY-CDLAHSHMQSIIDKLQSPESF 327  
855 GDRVVECOLETHNRKMTFKFOLD-GDNPEEATIMVNDFFILATERESFVDVREIEK 913  
328 AKSVQBELTIVLQRTGDPANLRLRPHLELANID-----PNPDVSPW 371  
914 ADEMUSEDSVEPEGQOG-----LESLOGKDDYGFSGQKLEGEFKQIOPASSMP-- 963  
372 EOLENAMAVKTVVHGLVDFIQNSRKGHETPOQPN-----SKYKTSMCRDLRQOG 423  
964 QQIGITSSILTVVH-----SAGRRFIVSPVPESRLRESKVPFSEITDVTAASTAQP 1016  
424 GCPRGNTCTFANSQOELEKVR-----LRNKKI-----NATVTEP-----LLNKV 463  
1017 G-----MNLSSHASSLSLQAFSELRAQMTGEPNTAPNFSHTGTFPVVPPFLSSIA 1070  
464 GVNNTVTFTA-----GNVISVIGSTETTKIVPSTNGISNAENS-----502  
1071 GVPTTAANTAPVPATSPNDISTSVIQSEVT---VPTTEGIAGVATSGVTVSGGLPIP 1127  
503 -----VSQISRSTDLRALETVKYKVGKAN-----QONAAGPSAD-SVTENKIG 548  
1128 PVSESPVLSVVSSITIPAVVSIITSPSLQVPTSTSEIWSSTALYPSVTVSATASAG 1187  
549 -----SPKPTPVSNVATSGPSNVGTSLNSV-----POKSSPELTV 586  
1188 GSTATGPKPPAVVSOQAAGSTTVGATLTSVSTTSPSTASQLSLSSSTSTPTLAET 1247  
587 PVYPHSENIQYFQDPTQIPREVQYPTQYGYPPPP-----TVRAGVAPCV-PRFVRSN 640  
1248 VVVAHSLD-KTSHSSTTGLAFSL-SAPSSSSFGAGVSYSIQPGGLHPLVIPSVIAT 1305  
641 NVPESSLPPASMEFYADHYSTFSPRDRMNSPYQPPPPQYGPVPPVPSGM-----690  
1306 PILPQAAGPTSTPL-----LPQVPSIPPLVQPVANVPAVQOTLIHSQPOPALL 1353  
691 -----YAPVYDSRRIRWPPMYORDDI-----IRS--NSLPPMDVWHSSV-YQTSRLR 735  
1354 PNQPHTHCPVDSDT--QPKAPGIDDIKTLEKLRSLFSEHSSSGAQSASVLSLETSLVIE 1411  
736 YNSLDGYYSVA-----COPPSE-PRTTVPLPREPC---GHLKTSCEEQIRR- 777  
1412 STVTGPIPTTAVAPSKLLTSTTCLPPTNLPLGTVALPVTVPVQVSTPVSTTSGV 1471  
778 KPDQWQYHTQKAPLVSS-TLPVATQSPPTPPLFSVDPRADPSESVSQTKFEEDHLSHY 836  
1472 KPQT-----APSKPLTKAPVLPVGTLP-----1495  
837 SPWSCGTIGSCINAIIDSEPKDVIANAVLMDLDS-----GDVKKRVH 879  
1496 -----AGTLPS--EQLPPFGPSLTQSQOPLDLDLAQLRRTLSPXITVTISAVGVPVMAAP 1549

QY 880 LFTQRRTK-----BEDPIIPSDGPIISKWG-----AISRSRTGYHTTDPVOAT 925  
Db 1550 TAITAAGTQPKQGVSVQKGVFVLTSSGAGVFKMRGQVSVAAADGAQKGNKSEDAKSV 1609  
QY 926 ASQGSATKPIVSVDVYVNAV-----DSRWSSYGNENATS-SAHYVER 967  
Db 1610 HFESSTSESSVSSSSPESTLVKPEPNCITIPGISSDVPESAHTTASEAKSDTQPTKV 1669  
QY 968 DRFIVTDLS---GHRKHSSTGDLIS-----LELOQAKSNLLLOREANALAM 1011  
Db 1670 GRFOVTTTANKVGRPSVSKTEDKITDTKKEGVPASPPFMDLEQAVLPAVIPKKEPELSE 1729  
QY 1012 QOKWN-----SLDEGR---HLTNLLSKEIELRN-----GELOSDY-TEDA 1048  
Db 1730 PSHLNGPSSDPEAAFLSRDVEDDGGSGSPHQLSSKSLPSONLSQSLSNSFNSSTMSSDN 1789  
QY 1049 TDTKPDRIEELSALDTPDPGQSEPIEELIDIOLGISSONDQLLN-----GMAVEN 1101  
Db 1790 ESDIEDEDLKLELRL-----RDHLKEIQDLSRQKHEIESLYTKLGKVPVAVIIPP 1842  
QY 1102 GHPVQOQHKEPPKQK-----KQSLGEDH-----VILEEQKTLIPVTSQFSG 1142  
Db 1843 AAPLSGRRRPTKSGKSSRSSSLGNKSPQLSGNLSGSAASVLPQOTLHPGPN---- 1898  
QY 1143 PLPVSIASCLPITTSVSAGNL 1165  
Db 1899 -IPESGQNLQPLKPPSSDNL 1920

Search completed: August 4, 2005, 01:49:10  
Job time : 56 secs

This Page Blank

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 01:27:38 ; Search time 192 Seconds

(without alignments)  
3176.491 Million cell updates/sec

Title: US-10-619-992-2

Perfect score: 6215

Sequence: 1 MPVQAQWTEFLSCPICYN.....MSDKNDFLKFVANGKQVNS 1191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Match | Length | ID | Description |
|------------|--------|-------|--------|----|-------------|
| 1          | 5483   | 78.2  | 1048   | 2  | Q9HBD2      |
| 2          | 4412   | 81.0  | 869    | 2  | Q9NKE1      |
| 3          | 3721   | 59.9  | 728    | 2  | Q8N3D6      |
| 4          | 2847   | 45.8  | 540    | 2  | Q9H5J2      |
| 5          | 2685   | 43.2  | 1136   | 2  | Q69231      |
| 6          | 2577.5 | 41.5  | 1109   | 2  | Q81VE6      |
| 7          | 2477   | 39.9  | 1114   | 2  | Q6NUC6      |
| 8          | 2293.5 | 36.9  | 506    | 2  | Q86ST6      |
| 9          | 2165   | 34.8  | 419    | 2  | Q9HBD1      |
| 10         | 1538   | 24.7  | 819    | 2  | Q9VV48      |
| 11         | 1289.5 | 20.7  | 398    | 2  | Q7PTD3      |
| 12         | 899    | 14.3  | 198    | 2  | Q9NWN9      |
| 13         | 843    | 13.6  | 1014   | 2  | Q45962      |
| 14         | 780    | 12.6  | 177    | 2  | Q96F27      |
| 15         | 611    | 9.8   | 522    | 2  | Q8N9V1      |
| 16         | 213.5  | 3.4   | 1157   | 1  | BBC1_YEAST  |
| 17         | 212    | 3.4   | 5085   | 1  | PCLO_RAT    |
| 18         | 211    | 3.4   | 1918   | 2  | Q752A6      |
| 19         | 207.5  | 3.3   | 1090   | 2  | Q9P5J0      |
| 20         | 207.5  | 3.3   | 1100   | 2  | Q7R2C6      |
| 21         | 206    | 3.3   | 1289   | 2  | Q9FLQ7      |
| 22         | 206    | 3.3   | 1611   | 2  | Q42854      |
| 23         | 205    | 3.3   | 1730   | 2  | Q6FJN8      |
| 24         | 204.5  | 3.3   | 9234   | 2  | Q7KTP5      |
| 25         | 204    | 3.3   | 2377   | 1  | WNK1_MOUSE  |
| 26         | 203    | 3.3   | 786    | 2  | Q48909      |
| 27         | 202    | 3.3   | 760    | 2  | Q9T0K5      |
| 28         | 201    | 3.2   | 1082   | 2  | Q8VIG1      |
| 29         | 200.5  | 3.2   | 728    | 2  | Q9HFN3      |
| 30         | 200    | 3.2   | 1134   | 2  | Q632W5      |
| 31         | 198    | 3.2   | 684    | 2  | Q6VQR0      |

Q6fay6 candida gla  
Q9h4a3 homo sapien  
Q97005 leishmania  
Q65375 arabidopsis  
Q78710 neurospora  
Q8cf91 mus musculu  
Q8cf92 mus musculu  
Q16936 caenorhabdi  
Q60885 homo sapien  
Q9gyx7 mus musculu  
Q80tw6 mus musculu  
Q8ia13 dictyosteli  
Q43432 homo sapien  
Q09085 phaseolus v

## ALIGNMENTS

### RESULT 1

Q9HBD2  
ID Q9HBD2 PRELIMINARY; PRT; 1048 AA.  
AC Q9HBD2;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Membrane-associated nucleic acid binding protein (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20517928; PubMed=10938276; DOI=10.1074/jbc.M004461200;  
RA Siess D.C., Veddor C.T., Merkens L.S., Tanaka T., Freed A.C.,  
RA McCoy S.L., Heinrich M.C., Deffebach M.E., Bennett R.M.,  
RA Hefenelder S.H.;  
RT "A human gene coding for a membrane-associated nucleic acid-binding protein."  
RL J. Biol. Chem. 275:33655-33662(2000).  
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
DR EMBL; AF255303; AKG00432.1; -;  
DR HSSP; P38398; 1JM7.  
DR GO; GO:0009986; C:cell surface; IDA.  
DR GO; GO:0005624; C:membrane fraction; IDA.  
DR GO; GO:0003677; F:DNA binding; IDA.  
DR InterPro; IPR000571; Znf\_CCCH.  
DR InterPro; IPR001841; Znf\_RING.  
DR Pfam; PF00642; zf-CCHC4; 1.  
DR Pfam; PF00642; zf-CCCH; 1.  
DR SMART; SM00184; RING; 1.  
DR SMART; SM00356; Znf\_C3H1; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; UNKNOWN\_1.  
DR PROSITE; PS00889; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
FT NON TER 1048 1048  
SQ SEQUENCE 1048 AA; 116077 MW; 0487D5B5C17D20C6 CRC64;

Query Match 88.2%; Score 5483; DB 2; Length 1048;  
Best Local Similarity 100.0%; Pred. No. 1.7e-265;  
Matches 1048; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPVQAQWTEFLSCPICYNFENVHKPISLGCSHTVCKTCLNKLHRKACPFDTAINTD 60  
DB 1 MPVQAQWTEFLSCPICYNFENVHKPISLGCSHTVCKTCLNKLHRKACPFDTAINTD 60  
QY 61 IDVLVFNFALLQVGAQVDPHQSIKLSNLGENKHVEVAKKVEDLALYKPLSGGKGVAS 120  
DB 61 IDVLVFNFALLQVGAQVDPHQSIKLSNLGENKHVEVAKKVEDLALYKPLSGGKGVAS 120  
QY 121 LQNSALSRRPMQKRLVTLVNCQLVEEGRVAMRAARSLGERTVTLLIHOHQNPQOLSANL 180  
DB 121 LQNSALSRRPMQKRLVTLVNCQLVEEGRVAMRAARSLGERTVTLLIHOHQNPQOLSANL 180

|                       |   |                                |     |
|-----------------------|---|--------------------------------|-----|
| OX                    | NCBI_TaxID=9606;  |                                |     |
| RN                    | [1]   |                                |     |
| RP                    | SEQUENCE FROM N.A.  |                                |     |
| RA                    | Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., |                                |     |
| RA                    | Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,  |                                |     |
| RA                    | Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isegai T., Sugano S.; |                                |     |
| RL                    | Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.               |                                |     |
| DR                    | EMBL; AK000308; BAA91073.1; --  |                                |     |
| DR                    | Genew; HGNC:8212; ORIK1.  |                                |     |
| DR                    | GO; GO:0003676; F:nucleic acid binding; IEA.                          |                                |     |
| DR                    | InterPro; IPR000571; Znf_CCH.   |                                |     |
| DR                    | Pfam; PF00642; zf-CCH; 1.   |                                |     |
| DR                    | SMART; SM00356; Znf_C3H1; 1.  |                                |     |
| SQ                    | SEQUENCE 869 AA; 96347 MW; 28105761C60B3D65 CRC64;                    |                                |     |
| Query Match           | 71.0%; Score 4412; DB 2; Length 869;                                  |                                |     |
| Best Local Similarity | 99.6%; Pred. No. 4.2e-212;  |                                |     |
| Matches               | 843; Conservative 2; Mismatches 1; Indels 0; Gaps 0;                  |                                |     |
| QY                    | 196 MOEEALKVLVLALEDGSALS  | RKVLVLFVVRORLEPRFPQASKTSTIGHVQ | 255 |
| Db                    | 1 MQEEALKVLVLALEDGSALS  | RKVLVLFVVRORLEPRFPQASKTSTIGHVQ | 60  |
| QY                    | 256 KRDESSLMQLKEEPRSYEARLRREHDAQIVHIAMEAGLRISPEQWSSLLYGLDIAHKSHMQ     | 315                            |     |
| Db                    | 61 KRDESSLMQLKEEPRSYEARLRREHDAQIVHIAMEAGLRISPEQWSSLLYGLDIAHKSHMQ      | 120                            |     |
| QY                    | 316 SIIDKLOSPESPAKSVQELTIVLQRTGDPANLRLRPHLELLANIDNPDAVSPTEWQLE        | 375                            |     |
| Db                    | 121 SIIDKLOSPESPAKSVQELTIVLQRTGDPANLRLRPHLELLANIDNPDAVSPTEWQLE        | 180                            |     |
| QY                    | 376 NAMVAVKTVVHGLVDFIQNYSRKGHETPQOPNSKYKTS                            | 435                            |     |
| Db                    | 181 NAMVAVKTVVHGLVDFIQNYSRKGHETPQOPNSKYKTS                            | 240                            |     |
| QY                    | 436 SOEELEKYLRLNKKINATVTFPLLNKVGNNVTVTTAGNVISVIGSTETTKGIVPSTNG        | 495                            |     |
| Db                    | 241 SOEELEKYLRLNKKINATVTFPLLNKVGNNVTVTTAGNVISVIGSTETTKGIVPSTNG        | 300                            |     |
| QY                    | 496 ISNAENSVQLISRSSTDSTLRALETVKVKGVKVGANGQNAAGPSADSVTENKIGSPKTPV      | 555                            |     |
| Db                    | 301 ISNAENSVQLISRSSTDSTLRALETVKVKGVKVGANGQNAAGPSADSVTENKIGSPKTPV      | 360                            |     |
| QY                    | 556 SNVAATAGPSNVGTSLNSVPQKSSPFLTRVPVYPHSENIQYFQDPRTQIPFVPOYPQ         | 615                            |     |
| Db                    | 361 SNVAATAGPSNVGTSLNSVPQKSSPFLTRVPVYPHSENIQYFQDPRTQIPFVPOYPQ         | 420                            |     |
| QY                    | 616 TGYPPPTVPAGVAPCVPRFVRNSNVPESSLPASMPYADHYSTFSPDRMNSSPYQPP          | 675                            |     |
| Db                    | 421 TGYPPPTVPAGVAPCVPRFVRNSNVPESSLPASMPYADHYSTFSPDRMNSSPYQPP          | 480                            |     |
| QY                    | 676 PPQYGPVPPVPSGMYAPVYDSRRIRWPPMYQRDDIIRNSLPMQVHSSVYQTS              | 735                            |     |
| Db                    | 481 PPQYGPVPPVPSGMYAPVYDSRRIRWPPMYQRDDIIRNSLPMQVHSSVYQTS              | 540                            |     |
| QY                    | 736 YNSLDGYSVACQPPSEPRTTVPLPREPCGHLKTSCEEQIRRKPDQWQYHTQKAPLVSS        | 795                            |     |
| Db                    | 541 YNSLDGYSVACQPPSEPRTTVPLPREPCGHLKTSCEEQIRRKPDQWQYHTQKAPLVSS        | 600                            |     |
| QY                    | 796 TLPVATQSPPTPPLFSVDFRADFSVSGTKFEDHLSHYSPWCGTIGSCINAIDSEP           | 855                            |     |
| Db                    | 601 TLPVATQSPPTPPLFSVDFRADFSVSGTKFEDHLSHYSPWCGTIGSCINAIDSEP           | 660                            |     |
| QY                    | 856 KDVIANNAVLMDLSDGDKRRVHLFETQRTKKEEDPIIPFSDGPIISKWGAISRSRTG         | 915                            |     |
| Db                    | 661 KDVIANNAVLMDLSDGDKRRVHLFETQRTKKEEDPIIPFSDGPIISKWGAISRSRTG         | 720                            |     |
| QY                    | 916 YHTTDPQATASOGSATKPIVSVDYVPVNAVDSRWSSYGNATSAAHYVERDRFIVTDL         | 975                            |     |
| Db                    | 721 YHTTDPQATASOGSATKPIVSVDYVPVNAVDSRWSSYGNATSAAHYVERDRFIVTDL         | 780                            |     |
| QY                    | 976 SGHRKHSSTGDDLLELQQAQKSNLLQREANALAMQKQWNSLDEGRHLTLNLLSKEIEL        | 1035                           |     |



Db 781 SGHRKHSSTGDLSSLSLQAKSNLSLLQREANALAMQKWNLSLDEGRHLTLNLLSKBIEL 840  
QY 1036 RINGELQ 1041  
| | | | |  
Db 841 RINGEVK 846  
| | | | |  
RESULT 3  
Q8N3D6 PRELIMINARY; PRT; 728 AA.  
AC Q8N3D6;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DE Hypothetical protein DKFZp762N035 (Fragment)  
GN Names=DKFZp762N035;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Melanoma;  
RG The German cDNA Consortium;  
RA Bloecher H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,  
RA Osanger A., Fobo G., Han M., Wiemann S.;  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
RE EMBL; AL834431; CAD39091.1; -  
DR GO; GO:0009986; C:cell surface; ISS.  
DR GO; GO:0005624; C:membrane fraction; ISS.  
DR GO; GO:0003677; F:DNA binding; ISS.  
DR InterPro; IPR000571; Znf CCCH.  
DR Pfam; PF00642; zf-CCCH; 1.  
DR SMART; SM00356; Znf C3H1; 1.  
KW Hypothetical protein.  
FT NON\_TER 1  
SQ SEQUENCE 728 AA; 80284 MW; 5B2B6C5682AA93A6 CRC64;  
Query Match 59.9%; Score 3721; DB 2; Length 728;  
Best Local Similarity 99.7%; Pred. No. 1.1e-177;  
Matches 703; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 337 VLQRTGDPANLRPHLELLANIDPNDAVSTWQLENAMVAVKTVVHGLVDFIQNY 396  
| | | | |  
Db 1 VLQRTGDPANLRPHLELLANIDPNDAVSTWQLENAMVAVKTVVHGLVDFIQNY 60  
| | | | |  
QY 397 RKGHETPQPNISKYKTSKCRDLRQGGCPRGTCNCTFAHSQBLEKYLRLNKKINATVRT 456  
| | | | |  
Db 61 RKGHETPQPNISKYKTSKCRDLRQGGCPRGTCNCTFAHSQBLEKYLRLNKKINATVRT 120  
| | | | |  
QY 457 FPLLNKVGNNVTVTAGNVISVIGSTETTKIVPSTNGISNAENSVSOLISRSTDTLRL 516  
| | | | |  
Db 121 FPLLNKVGNNVTVTAGNVISVIGSTETTKIVPSTNGISNAENSVSOLISRSTDTLRL 180  
| | | | |  
QY 517 ALETVKVGKVGANGQNAAGPSADSVTENKIGSPKTPVSNVAATSAAGSNVGTSLNSVP 576  
| | | | |  
Db 181 ALETVKVGKVGANGQNAAGPSADSVTENKIGSPKTPVSNVAATSAAGSNVGTSLNSVP 240  
| | | | |  
QY 577 QKSSPELTVVPVPPHSENIQYQDPRTOIPEFVPOYPQTGYPPPTPVAGVAPCVPRP 636  
| | | | |  
Db 241 QKSSPELTVVPVPPHSENIQYQDPRTOIPEFVPOYPQTGYPPPTPVAGVAPCVPRP 300  
| | | | |  
QY 637 VRNNNPESLPPASMPYADHYTFSPDRMNSSPYQPPPPQYGPVPPVPSPGMYAPVVD 696  
| | | | |  
Db 301 VRNNNPESLPPASMPYADHYTFSPDRMNSSPYQPPPPQYGPVPPVPSPGMYAPVVD 360  
| | | | |  
QY 697 SRRIWRPMPYQRDDIIRNSNLPMDVMHSSVYQTSRLRYNSLDGYYSVACQPPSEPRIT 756  
| | | | |  
Db 361 SRRIWRPMPYQRDDIIRNSNLPMDVMHSSVYQTSRLRYNSLDGYYSVACQPPSEPRIT 420  
| | | | |  
QY 757 VPLPREPCGHLKTSCEEQIRRKPDQWQAQYHTQKAPLVSTLPVATQSPTPPSPLFSDVFR 816  
| | | | |  
Db 421 VPLPREPCGHLKTSCEEQIRRKPDQWQAQYHTQKAPLVSTLPVATQSPTPPSPLFSDVFR 480  
| | | | |

QY 817 ADFSESVSCTKFEEDHLSHYSPWSCGTIGSCINAIDSEPKDVIANNAVLMOLDSDGVKR 876  
| | | | |  
Db 481 ADFSESVSCTKFEEDHLSHYSPWSCGTIGSCINAIDSEPKDVIANNAVLMOLDSDGVKR 540  
| | | | |  
QY 877 RVHLFETQRTKEEDPIIPFSDGPIISKWGAISRSTGYHTTDPVOATASQGSATKPLS 936  
| | | | |  
Db 541 RVHLFETQRTKEEDPIIPFSDGPIISKWGAISRSTGYHTTDPVOATASQGSATKPLS 600  
| | | | |  
QY 937 VSDYVPYVNAVDSRWSSYNGEATSSAHYVERDRFIVTDLSGHRKHSSTGDLSSLELQOAK 996  
| | | | |  
Db 601 VSDYVPYVNAVDSRWSSYNGEATSSAHYVERDRFIVTDLSGHRKHSSTGDLSSLELQOAK 660  
| | | | |  
QY 997 SNSLLQREANALAMQKWNLSLDEGRHLTLNLLSKBIELRNGELQ 1041  
| | | | |  
Db 661 SNSLLQREANALAMQKWNLSLDEGRHLTLNLLSKBIELRNGEVK 705  
| | | | |  
RESULT 4  
Q9H5J2 PRELIMINARY; PRT; 540 AA.  
AC Q9H5J2;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein FLJ23389.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,  
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK027042; BAB15634.1; -  
SQ SEQUENCE 540 AA; 60158 MW; 8B040AFB87F17A27 CRC64;  
Query Match 45.8%; Score 2847; DB 2; Length 540;  
Best Local Similarity 100.0%; Pred. No. 3.3e-134;  
Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 652 MPYADHYSTFSPDRMNSSPYQPPPPQYPGPVPPVPSGMYPYVDSRRIRWRPMPYQDDI 711  
| | | | |  
Db 1 MPYADHYSTFSPDRMNSSPYQPPPPQYPGPVPPVPSGMYPYVDSRRIRWRPMPYQDDI 60  
| | | | |  
QY 712 IRSNSLPPMDVMHSSVYQTSRLRYNSLDGYYSVACQPPSEPRITVPLPREPCGHLKTS 771  
| | | | |  
Db 61 IRSNSLPPMDVMHSSVYQTSRLRYNSLDGYYSVACQPPSEPRITVPLPREPCGHLKTS 120  
| | | | |  
QY 772 EQIIRRKPDQWQAQYHTQKAPLVSTLPVATQSPTPPSPLFSDVFRADFSESVSCTKFEED 831  
| | | | |  
Db 121 EQIIRRKPDQWQAQYHTQKAPLVSTLPVATQSPTPPSPLFSDVFRADFSESVSCTKFEED 180  
| | | | |  
QY 832 HLSHSPWSCGTIGSCINAIDSEPKDVIANNAVLMOLDSDGVKRRVHLFETQRTKEED 891  
| | | | |  
Db 181 HLSHSPWSCGTIGSCINAIDSEPKDVIANNAVLMOLDSDGVKRRVHLFETQRTKEED 240  
| | | | |  
QY 892 PIIPFSDGPIISKWGAISRSTGYHTTDPVOATASQGSATKPLISVSDYVPYVNAVDSRW 951  
| | | | |  
Db 241 PIIPFSDGPIISKWGAISRSTGYHTTDPVOATASQGSATKPLISVSDYVPYVNAVDSRW 300  
| | | | |  
QY 952 SSYNGEATSSAHYVERDRFIVTDLSGHRKHSSTGDLSSLELQAKSNLSLLQREANALAM 1011  
| | | | |  
Db 301 SSYNGEATSSAHYVERDRFIVTDLSGHRKHSSTGDLSSLELQAKSNLSLLQREANALAM 360  
| | | | |  
QY 1012 QKWNLSLDEGRHLTLNLLSKBIELRNGELQSDYTEDATDTKPRDRIEELSALDTDEPDG 1071  
| | | | |  
Db 361 QKWNLSLDEGRHLTLNLLSKBIELRNGELQSDYTEDATDTKPRDRIEELSALDTDEPDG 420  
| | | | |  
QY 1072 QSEPIEEIIDIQLGISNDQQLNGMAVENGHVPQHQKEPPKQKQSGEDHVLIEBQK 1131  
| | | | |

Db 421 QSEPIEELIDQLGISSQNDQLNGMAVENGHVPQHQKEPPKQKQSLGDHVLBEQK 480  
 QY 1132 TLTPVTSFQSLPVPISNASCLPITTSVAGNLILKTHVMSDQNDFLKPVANGKMYNS 1191  
 Db 481 TLTPVTSFQSLPVPISNASCLPITTSVAGNLILKTHVMSDQNDFLKPVANGKMYNS 540  
 RESULT 5  
 Q69231  
 ID Q69231 PRELIMINARY; PRT; 1136 AA.  
 AC Q69231;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE MKIAA2025 protein (Fragment).  
 GN Name=mkIAA2025;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,  
 RA Saga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,  
 RA Nagase T., Ohara O., Koga H.;  
 RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:  
 IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous  
 RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones  
 RT Randomly Sampled from Size-Fractionated Libraries.";  
 RL DNA Res. 11:205-218(2004).  
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
 DR EMBL; AK173335; BAD32613.1; -  
 DR GO; GO:000151; C:ubiquitin ligase complex; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:001567; P:protein ubiquitination; IEA.  
 DR InterPro; IPR000571; Znf\_CCH.  
 DR InterPro; IPR001841; Znf\_Ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR Pfam; PF00642; zf-CCCH; 1.  
 DR SMART; SM00184; RING; 1.  
 DR SMART; SM00356; Znf\_C3H1; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS00589; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON TER 1  
 SQ SEQUENCE 1136 AA; 125988 MW; 4A3A2BA6F13C0D51 CRC64;

Query Match 43.2%; Score 2685; DB 2; Length 1136;  
 Best Local Similarity 52.2%; Pred. No. 1.2e-125;  
 Matches 615; Conservative 135; Mismatches 295; Indels 134; Gaps 36;

QY 1 MPVQAQWTEFLSCPICYNEFDNVHKPISLGCSTHVTCKLNLKLRKACPDPQTINTD 60  
 Db 7 MPVQAQWTEFLSCPICTQTFDETIRKPSLGCSTHVTCKLNLKLRKACPDPQTINTD 66  
 QY 61 IDVLVFNFALLQVLGVAQVPHQSIKL-SNLGENKHVEYAKKCVEDIALYLKPLSGKGVA 119  
 Db 67 IELLVFNFALLQVLGVAQVPHQSIKL-SNLGENKHVEYAKKCVEDIALYLKPLSGKGVA 126  
 QY 120 --SLNQSALSRPQKLVTLVNCQLVEERGRVAMPAARSLGERTVTELILOHQNPOOLS 177  
 Db 127 LNSTTQSVLSRPQKLVTLVNCQLVEERGRVAMPAARSLGERTVTELILOHQNPOOLS 186  
 QY 178 ANLWAAVRAGCCFLGPAQMEERALKVLVLALEDGSLSRKVLVFLVQRLPRFPQASKT 237  
 Db 187 SNLWAAVRAGCCFLGPAQMEERALKVLVLALEDGSLSRKVLVFLVQRLPRFPQASKT 246  
 QY 238 SIGHVVQLLYRASCFTVKTRDESSLQMLKEEFYSYEAALRRHDAQIVHAMEAGLRISP 297  
 Db 247 SIGHVVQLLYRASCFTVKTRDESSLQMLKEEFYSYEAALRRHDSQIVQIAMEAGLRISP 306

QY 298 EQWSSLYGDLAKSHMQSIIDKLQSPESFAKSVQELTIVLQRTGDPANLRLRPHLELL 357  
 Db 307 DQWSSLYGDLAKSHMQSIIDKLQSPASFAQSVELTIALQRTGDPANLRLRPHLELL 366  
 QY 358 ANIDPNDAVPTWEQLENAMAVKTVVHGLVDFIYQNSRKGHETPPQOPNSKYKTSMCR 417  
 Db 367 ANIDPSPDAPPTWEQLENGLVAVRTVVHGLVDFIYQNSRKGADQQQPPQHSKYKTYMCR 426  
 QY 418 DLROGGCGPRTNCTFAHSELEKYLRLNKKINATVTFPLLNKGVNNTVTAGNVI 477  
 Db 427 DMKRGCGPRTNCTFAHSELEKYLRLNKKINATVTFPLLNKGVNNTVTAGNVI 483  
 QY 478 SVIGSTETTGKIVPS-TNGISNAENSQSLISRSTDL-RALETVKVKGVKGVANGQNA 535  
 Db 484 SDESVDLSNRKPPALPNGIASSGSTVTLIPRGTDPSDFSLSKPVK----- 530  
 QY 536 GPSADSVTENKIGSPKTPVSNVAATSAGPNVGTLSNVQKSPFFLTPVTV----- 588  
 Db 531 ---LDHLSSAPGSPDLSAPKSIAPVNP-----HPVPPRGPTDLPMPVTKPIQMV 583  
 QY 589 -----YPPHSENIOYFODPRTQIP-FEVPOYPOGYPPPPPTVPAGVAPCVPRVRSNN 641  
 Db 584 PRGSOLYPAQADV--YYQDPGRGSAFAFETAPYQQGMYYTPP-----PCVSRFVR--- 631  
 QY 642 VPESLPPASMPYADHYSTFSPDRMNSSPYQPPPPQPYGPPVPSGMYAPVYDSRIW 701  
 Db 632 -PPPSAPFPGPYLDHYSPLY-QDRVINSYGTQPPQ---YPP---MYPAHYDGRVY 681  
 QY 702 RPPMYQRDDIIRNSLPPMDVHMSV--YQTSRLRYNSLDGYYVACOP-----PSEPR 755  
 Db 682 PAQSYTREEMFRESPI-DIDIPSAAPVSVESRYQYVEGYVPVAPHPAQRFSYPRD 740  
 QY 756 TVPLPREPCGHLKTSCEEQIRKPPQWQAYHTOKAPLVSTLTPVATQSPPTPSPLFSVDF 815  
 Db 741 P-PYSRLPPPPQPHPSLDLHRRRKEIMAELEERK---VISPPFFA-PSPTLP-PAFHPBE 794  
 QY 816 RADFSESYSVGTKEEDHLSHYSPWSCGTIGSCINAIIDSEPKDVIANNAVLMDLDS-GDV 874  
 Db 795 FLDEDLKVAG-KYKANDYSQYSPNSCDTIGSYIGTKDAKPKDVVAAGSVEMNANVESKGR 853  
 QY 875 KRVHLFTQRTKE--EDPIPPSDGPIISKWGAISRSTGYHTTDPVQATASGASAT 932  
 Db 854 EQRLDL---QRRAVETSDDLIPFGDRPTVSRFGAISRTSKTYQAGLQAIAPQAGPT 910  
 QY 933 KPIVSVDYVPY---VNAVDSRWSSYNEATSSAHVVERDRFTVTDLSGHRKH--SSTGL 987  
 Db 911 KSNISDISAYGAHGGWGDSPYSFPHAN-IPPOGHFIEREKMSMAEVAHSGKLLSAERQ 969  
 QY 988 LSLELQQ-----AKNSLLQLQREANALAMQ-----QKWSLDEGRHLT 1025  
 Db 970 LLELQLNLHQLSQQLRGLAEVSNRLVQLREVNTLASQPPQLPPKWPGMISSEQLS 1029  
 QY 1026 LNL--LSKEIENLNGELSDYTEDATDKPRDRIELELSALDTDE--PDQSEPTIEILD 1081  
 Db 1030 LELHQVEREIGKRTRELSME-NQCSVDMK-----SKLGTSKQAENGQPEPQNKI-- 1077  
 QY 1082 IQGLISSQNDQLNGMAVENGHVPQHQKEPPKQKQSL 1120  
 Db 1078 -----RTEDLTLTFTSDVPNGSALTQENLSLNSKTSLSL 1110

## RESULT 6

Q8IVE6 PRELIMINARY; PRT; 1109 AA.  
 AC Q8IVE6;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE KIAA2025 protein (Fragment).  
 GN Name=KIAA2025;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Nagase T., Kikuno R., Ohara O.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB095945; BAC23121.1; -.
DR HSP; P22893; 1M90.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR000571; Znf_CCH.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00642; zf_CCH; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
KW Metal-binding; Zinc; Zinc-finger.
FT NON TER 1
SQ SEQUENCE 1109 AA; 122994 MW; 1D0D6A5F3E1497EC CRC64;

Query Match 41.5%; Score 2577.5; DB 2; Length 1109;
Best Local Similarity 51.3%; Pred. No. 2.7e-120;
Matches 595; Conservative 135; Mismatches 290; Indels 139; Gaps 37;

QY 25 VHKPISLGSCHTVCCKLKLKHKACPFDPQTATNDIDVLPVNFALLQLVGAQVDPHQSI 84
DB 1 IRKPISLGCCHTVCKMLKHLKHKACPFDPQTINTDIELLPVNSALLQLVGAQVPEQPI 60

QY 85 KL-SNLGENKHYEAKKVEDLALYLKPLSGGKGA--SLNQSALSRRMQRKLVTLVNCQ 141
DB 61 TLCSGVEDTKHYEAKKVEELALYLKPLSSARGVGLNLTQSVLSRPMQRKLVTLVHCQ 120

QY 142 LVEEGRVPMARAARSLGRTVTTELILQHNQFQOLSANLWAAVRARGCOFLGPAHQEAL 201
DB 121 LVEEGRIPARAARSLGRTVTTELILQHNQFQOLSSNLWAAVRARGCOFLGPAHQEAL 180

QY 202 KLVLLALEDGSALSARKVLVFLVVRLEPRFPQASKTSIGHVQLLYRASCFFVKTRDEDS 261
DB 181 KLVLLALEDGSALSARKVLVFLVVRLEPRFPQASKTSIGHVQLLYRASCFFVKTRDEDS 240

QY 262 SLMQKKEEPRSYEARLREHDAQIVHIAMEAGLRISPEQSSLLYGDLAHKSQMHSIIDKL 321
DB 241 SLMQKKEEPRTEALRREHDSQIVQIAMEAGLRAPDQSSLLYGDQSKHMQHSIIDKL 300

QY 322 QSPESPAKSQVELTIVLQRTGDPANLRLRPHLELLANDPNPDVSPTEQLENAMVAV 381
DB 301 QTPASPAQSQVELTIALQRTGDPANLRLRPHLELLANDPSPDAPPPTWEQLENGLVAV 360

QY 382 KTVVHGLVDFIQNSRKGHETPOPNKSKYKTSMCRLDRQCGCPRGTCNCTFAHSQEELE 441
DB 361 RTVVHGLVDYIQHNSKKGADQQQPPQHSKITYKTMCKDMKQRCGCPKASCTFAHSQEELE 420

QY 442 KYRLRNKKINATVTRTPPLNKKVGNVNTVTTAGNVISVIGSTETGKIYPS-TNGISNAE 500
DB 421 KFRKNNKRL--VPRPLSASLQQLNEVGLPSAAIIPDGSADVLPDSRKKPALPNGIVSTG 477

QY 501 NVSQILISSTSTLRALETVKVKGKVGANGQNAAGPSADSVTENKIGSPPTKTPVSNVAA 560
DB 478 NTVTQIIPRGTD--SYDSSLKPGKI-----DHLSSAPGSPFPLLESVPKSS 522

QY 561 TSAGPSNVGTSLNSVPKSSPFLTRVPV-----YPPHSENIQYFQDPR-TQI 606
DB 523 ISALFVNP-----HSIIPRGADLPMPVTKPLQMPVRGSQLPAQQTQD-VYQDPKGAAP 577

QY 607 PFEVPOYPTGYPPPTPTVAGVAPCVPRFVRSNNVPSSLPASPADYADHSTFSPDR 666
DB 578 PEPAPYQQGMYYTPP-----QCVSRFVR-----PPPSAEPAPYLDHPPYL-QER 625

QY 667 MNSSPYQPPPPQYGPVPPVPSGMVAPVYDSRRIRMPMYQRDITRSNLSLPMVMHSS 726
DB 626 VVNSQYGTQPOQ-----YPII-----YPSHYDGRVAPVSYTREIIPRESPI-PIEIPPA 676

102 V--YOTSLRERYNSLDGYYSVACOP-----PS-----EPRTTVPLPREPCGHLKTSCEQIR 776
103 VPSVPESEYQIETSYYPVAPHPTQIRPSYLRPPYSLRPPPPQ-----HPSLDELHR 732
104 RKPDMQAQYHTQKAPLVSTLPAVATQSPSPPLFSVDPRADFSVSSTKFEEDHLSHY 836
105 RRKEIMAELEERK--VLSPPFFA-PSPTLP-PTPHPEFDEDELDKVAG-KYKGNDSYQY 786
106 SPWSCGTIGSCINAIIDSEPKDVIANSNAVLMDLSDGDKVRRVHLFETQRTKE--EDPII 894
107 SPWSCDTIGSYIGTKDAKPKDVAAGSVEMNVESKGM--RDQRLDLQRAAETSDDDLI 844
108 PFDGPIISKMGCAISRSRTGHTTDPVQATASQSAKPSVSDYVPVNAVDSRW--S 952
109 PFGDRPTVSRFGAISRTSKTIYQAGPMQAMAPQAGPTKSNISIDSPY--GTHGGWGAS 902
110 SYG--NEATSSAHYVERDFIVTDLSGHRK--HSSTGDLLSLELOQ----- 994
111 PYSHPQNTPSQGHFSERERISMSEVASHGKPLPSAERQRLLEQLLQNLHQISQQTQLRGL 962
112 AKNSLLQREANALAMQ-----KWSLDEGRHLTLNL--LSKEIELRNGELOSD 1043
113 EAVSNRLVLQREANTLACSQSPPPPPKMPCGMISSEQLSLELHQVREIGKRTRELSME 1022
114 YTEDATDTKPRDIELELSALOTDE--PDGQSEPIEEILDQLGSISSQNDQLNGMAVEN 1101
115 -NQSLEDMK-----SKLNTSKQAEQGPQONKV-----PAEDLTFTSDVPN 1064
116 GHPVQOQKEPPKQKQSL 1120
117 GSALTQENISLLSNKTS 1083

RESULT 7
Q6NUC6 PRELIMINARY; PRT; 1114 AA.
AC Q6NUC6;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE MGC81061 protein.
GN Name=MGC81061;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues Y., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
```

RC TISSUE=Ovary;  
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.,  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative";  
RL Dev. Dyn. 225:384-391(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RA Klein S., Strausberg R.,  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.  
DR EMBL; BC068669; AAH66669.1; -  
DR GO; GO:0000153; Cubiquitin ligase complex; IEA.  
DR GO; GO:0003676; F-nucleic acid binding; IEA.  
DR GO; GO:0004842; F-ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0008270; F-zinc ion binding; IEA.  
DR GO; GO:0016567; P-protein ubiquitination; IEA.  
DR InterPro; IPR000571; Znf\_CCH.  
DR InterPro; IPR001841; Znf\_Ring.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR Pfam; PF00642; zf-CCHC; 1.  
DR SMART; SM00184; RING; 1.  
DR SMART; SM00356; Znf\_C3H1; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS00089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
SQ SEQUENCE 1114 AA; 123576 MW; E836C7514BF2A3B7 CRC64;  
  
Query Match 39.9%; Score 2477; DB 2; Length 1114;  
Best Local Similarity 47.3%; Pred. No. 2.8e-115;  
Matches 592; Conservative 140; Mismatches 310; Indels 210; Gaps 40;  
  
QY 1 MPVQAQMTFLSCPICYNFEDNVHVKPISLGCSTVCKTLNKLHAKACPFDPQAINVD 60  
DB 1 MPVQAQMTFLSCPICYQTFDTETTRKPISLGCGHTVCWCLNKLHAKACPFDPQITIND 60  
  
QY 61 IDVLVNFALLQVGAQVDPHOSIKL---SNLGENKHVEYAKKCVEDIALYKPLSGGKG 117  
DB 61 IELLFPVNSALLQVGAQVPEQQOITLGGCGAEDTKHYEARKCVELALYKPLSLTARG 120  
  
QY 118 VA--SLNQSLGRPMOKLVTLVNCOLVEEGRVRAARAARSIGERTVTTELILQHNPQQ 175  
DB 121 VGENSTQSVLRPMQKLVTLVHQLVVEEGRIRARAARSIGERTVTTELILQHNPQQ 180  
  
QY 176 LSNLWAAVRAQCQFLGPAMQEEALKVLLALEDGSALSRKVLVLFVVQRLEPRPQAS 235  
DB 181 LSNLWAAVRAQCQFLGPAMQEEALKVLLALEDGSALSRKVLVLFVVQRLEPRPQAS 240  
  
QY 236 KTSIGHVQVLLYRASCFTKTEDESSLMOLEKERSYEALREHDAQIVHTAMEAGLRI 295  
DB 241 KTSIGHVQVLLYRASCFTKTEDESSLMOLEKERTYEALREHDSQIVQVAMEAGLRI 300  
  
QY 296 SPEQWSSLLYGLAHKSHMQSIIDKLQSPSEFASKSVQELTIVLQRTGDPANLRLRPHLE 355  
DB 301 APDQWSSLLYGLQSHKSHMQSIIDKLQTPASFAQSQVQELTIALQRTGDPANLRLRPHLE 360  
  
QY 356 LLANIDPNPDVSPTEQLENAMAVKTVVHGLVDFIYONYSRKGHETPQPNSKYKTSM 415  
DB 361 LLANIDPSPDAPPTTEQQLKGLVAVKTVVHGLVDYIQNHKKGTQQQPQHSKYKTM 420  
  
QY 416 CRDLROQGCPRGNTCTFAHSGEELKYLRLNKKINATVTRPPL-----LNKVGNNVT 469  
DB 421 CRDMKQRCGPRGASCTFAHSGEELKFRMKNRL---VPRPLSASLQQLNEVGL----- 473  
  
QY 470 TTAGNVISGISTETTGKIVP--STNGISNA-----ENSVQSILSTSDTSLRALETVKV 524  
DB 474 -----PIGAPDEGMDLPRKPSGLPNGIVPGSSVTQLSRSDS---GFESVLKP 521  
  
QY 525 GKVGANGQNAAGSADSVTENKIGSPKTPVSNV-----AATSGAPNVGTGLNSVPQ 577  
DB 522 VKLDHLSSAPSPPE-----LLDSVPKTSMSALPVNAHPAASRDLPPLFVSKQIQWVR 576

## RESULT 8

Q86ST6  
ID Q86ST6 PRELIMINARY; PRT; 506 AA.  
AC Q86ST6;  
DT 01-JUN-2003 (TReMBLrel. 24, Created)  
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)  
DT 01-WAR-2004 (TReMBLrel. 26, Last annotation update)  
DE Hypothetical protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stachenko M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.L., Skalska U., Smallos D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Straussberg R.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
DR EMBL; BC044642; AAH44642.1; -;  
DR HSP; P22893; IM90.  
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0008270; F:zinc ion binding; IEA.  
DR GO; GO:0016567; P:protein ubiquitination; IEA.  
DR InterPro; IPR000571; Znf\_CCH.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR Pfam; PF00642; zf-CCCH; 1.  
DR SMART; SM00184; RING; 1.  
DR SMART; SM00356; Znf\_C3H1; 1; UNKNOWN\_1.  
DR PROSITE; PS00518; ZF\_RING\_1; UNKNOWN\_1.  
DR PROSITE; PS00889; ZF\_RING\_2; 1.  
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.  
SQ SEQUENCE 506 AA; 56853 MW; F2915DA0B8D0AC9E CRC64;

Query Match 36.9%; Score 2293.5; DB 2; Length 506;  
Best Local Similarity 95.5%; Pred. No. 1.4e-106;  
Matches 448; Conservative 3; Mismatches 13; Indels 5; Gaps 1;  
QY 1 MPVQAAQWTEFLSCPIYNEFDENVHKPISLGCSTHTVCKTLNKLHRCACPPDQTAINTD 60  
DB 1 MPVQAAQWTEFLSCPIYNEFDENVHKPISLGCSTHTVCKTLNKLHRCACPPDQTAINTD 60  
QY 61 IDVLPVNFALLQVGAQVPHQSIKLSNLGENKHVEAKKVEDLALYLKPLSGGKGVAS 120  
DB 61 IDVLPVNFALLQVGAQVPHQSIKLSNLGENKHVEAKKVEDLALYLKPLSGGKGVAS 120  
QY 121 LNQSALSRPMQKLVTLNQCQLVEEGRVRAARAARSLGERTVTELILOHQPQOLSANL 180  
DB 121 LNQSALSRPMQKLVTLNQCQLVEEGRVRAARAARSLGERTVTELILOHQPQOLSANL 180  
QY 181 WAAVRAGCQFLGPAMQEBALKVLALDGGALGRKVLVLFVQRLPRFPQAKTSIG 240  
DB 181 WAAVRAGCQFLGPAMQEBALKVLALDGGALGRKVLVLFVQRLPRFPQAKTSIG 240  
QY 241 HVVQLLYRASCFTVKTKRDESSLMQLKEFRSYEARLREHDAQIVHIAEAGLRISPEQW 300  
DB 241 HVVQLLYRASCFTVKTKRDESSLMQLKEFRSYEARLREHDAQIVHIAEAGLRISPEQW 300  
QY 301 SLLLYGDLAHKSHMQSIIDKLQSPESFAKSVQELTIIVLQRTGDPANLRLRPHLELLANI 360  
DB 301 SLLLYGDLAHKSHMQSIIDKLQSPESFAKSVQELTIIVLQRTGDPANLRLRPHLELLANI 360  
QY 361 DNPDAVSTPWELENAMAVKTVVHGLVDFTONYSRKGHTPOPOPNSKYKTSKCRDLR 420  
DB 361 DNPDAVSTPWELENAMAVKTVVHGLVDFTONYSRKGHTPOPOPNSKYKTSKCRDLR 420  
QY 421 QQGGCPRGNTCTFAHSEEBEELKYNLNKKNATVTRTFPLLNKVVGNNTV 469  
DB 421 QQGGCPRGNTCTFAHSEEBEELKYNLNKKNATVTRTFPLLNKVVGNNTV 469

RESULT 9  
Q9HBD1  
ID Q9HBD1

PRELIMINARY; PRT; 419 AA.

AC Q9HBD1;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE Membrane-associated nucleic acid binding protein (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20517928; PubMed=10938276; DOI=10.1074/jbc.M004461200;  
RA Siees D.C., Vedder C.T., Merkens L.S., Tanaka T., Fred A.C.,  
RA McCoy S.L., Heinrich M.C., Delfebach M.E., Bennett R.M.,  
RA Hefeneider S.H.;  
RT "A human gene coding for a membrane-associated nucleic acid-binding  
RT protein.";  
RL J. Biol. Chem. 275:33655-33662(2000).  
DR EMBL; AF255304; AAG00433.1; -;  
DR GO; GO:0009986; C:cell surface; IEA.  
DR GO; GO:0005624; C:membrane fraction; IEA.  
DR GO; GO:0003677; F:DNA binding; IEA.  
FT NON\_TER 1  
SQ SEQUENCE 419 AA; 46285 MW; 2C525A12C77F56DC CRC64;  
Query Match 34.8%; Score 2165; DB 2; Length 419;  
Best Local Similarity 100.0%; Pred. No. 2.7e-100;  
Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 773 EQIRKPDQWQAQYHTQKAPLVSSSTLPVATQSPPTPSPLFSDVDFRADFSVSGTKFEEDH 832  
DB 1 EQIRKPDQWQAQYHTQKAPLVSSSTLPVATQSPPTPSPLFSDVDFRADFSVSGTKFEEDH 60  
QY 833 LSHYSPWSCGTIGSCINAIDSEPKDVIANSNAVLMDLSDGVKRRVHLFETQRTKEEDP 892  
DB 61 LSHYSPWSCGTIGSCINAIDSEPKDVIANSNAVLMDLSDGVKRRVHLFETQRTKEEDP 120  
QY 893 IIPFSDGPIISKWGAISRSRTGYHTTDPVQATASQGSATKPIVSVDYVYVYNAVDSRWS 952  
DB 121 IIPFSDGPIISKWGAISRSRTGYHTTDPVQATASQGSATKPIVSVDYVYVYNAVDSRWS 180  
QY 953 SYGNEATSAHYVERDFIVTDLSGHRKHSSTGDLISLELOQAKNSLLLOREANALAMQ 1012  
DB 181 SYGNEATSAHYVERDFIVTDLSGHRKHSSTGDLISLELOQAKNSLLLOREANALAMQ 240  
QY 1013 QKNSLSDGRLHLLNLSKEIERNLNGELQSDYTEDATDKPRDIELELSALDTDEPDQ 1072  
DB 241 QKNSLSDGRLHLLNLSKEIERNLNGELQSDYTEDATDKPRDIELELSALDTDEPDQ 300  
QY 1073 SEPIBEILDIGLISSQNDQLNGMAVENGHVPVQHQKEPPKQKQSLGEDHVLLEEQKT 1132  
DB 301 SEPIBEILDIGLISSQNDQLNGMAVENGHVPVQHQKEPPKQKQSLGEDHVLLEEQKT 360  
QY 1133 ILPVTSCFQPLVPSISNASCLPITTSVAGNLIKTHVMSDKNDPLKPVANGKVVNS 1191  
DB 361 ILPVTSCFQPLVPSISNASCLPITTSVAGNLIKTHVMSDKNDPLKPVANGKVVNS 419  
RESULT 10  
Q9VV48  
ID Q9VV48 PRELIMINARY; PRT; 819 AA.  
AC Q9VV48;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE CG16807-PA.  
GN ORFNames=CG16807;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abril J.F., Agayani A., An H.J., Andrews-Frankoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster G., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hoston D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupezki M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasmann D.A., Weissstock G.M., Weissbach J.,  
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RT Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
RT melanogaster euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
RT a genomics perspective.";  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review.";

RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.  
DR EMBL; AE003527; AAF49474.1; -.  
DR FLYBase; FBgn0036621; CG16807.  
DR GO; GO:000151; C:ubiquitin ligase complex; IEA.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0008270; F:zinc ion binding; IEA.  
DR GO; GO:0016567; P:protein ubiquitination; IEA.  
DR InterPro; IPR001841; Znf CCH.  
DR Pfam; PF00097; Zf-C3HC4\_1.  
DR Pfam; PF00642; Zf-CCH; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS00089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
SQ SEQUENCE 819 AA; 90376 MW; 33D429076DE7E2E CRC64;  
Query Match 24.7%; Score 1538; DB 2; Length 819;  
Best Local Similarity 40.5%; Pred. No. 1.5e-68;  
Matches 386; Conservative 98; Mismatches 226; Indels 242; Gaps 29;  
QY 1 MPVQAQWTEFLSCPICYNFEEDNVHVKPISLGCSHTCKLNLKLRKACPFDDTAINTD 60  
DB 1 MPIQAPSWTDFELPCICCNFEAASQRCPSVLCGGHTICKLCTLYNQCFDQIVVSD 60  
QY 61 IDVLVFNALLQLV-----GAQVPDQHSIKLSNLGENKHYEYAKCEVEDLALYLKPL 112  
DB 61 IDNLPINHALQLVKDSLLELAPPPSVQKLEPHL---KCYQLGQRCIEELALHLKSF 117  
QY 113 SGGKGVASLNOSALSRRPMOKLVTLVNCQVVEEGRVRAARSLGERTVTELLIQHN 172  
DB 118 LNLNG-----NGNLLTRPMLRKLVTLVNCQVVEEGRVRAARSLGERTVTELLIQHN 173  
QY 173 POOLSANLWAAVRARGCOFLGPAMQEEALKLVLLALEDEGSALSRKVLVLFVVRLEPRFP 232  
DB 174 POOLSSNLWAAVRTRGCOFLGPAMQEEVLKLVLLALEEGSALSRLVWVFWVRLEPRFP 233  
QY 233 QASKTSIGHVVQLLYRASCYKVKTRKDESSLMQLKEEFSYVEALRREHDAQIVHIAMEAG 292  
DB 234 QASKTSIGHVVQLLYRASCYKVKREADSSLMQLKEEFTYDALRREHDAQIVQIATEAG 293  
QY 293 LRISPEQWSSLLYGDLAHSHMQSIIDKLOSPESPAKSVQELTVLQRTGDPANLRLRP 352  
DB 294 LRIAPEQWSSLLYGDVMEKSHMQSIIDKLOTPSSFAQSVQELVIALQRTSDPAKLSLHH 353  
QY 353 HLELANIDENPDVAVPTWEOLENAWAVKTVHGLVDFIQNYS-RK-----GHETPQ 404  
DB 354 HUKYLANIDPCAE-VAP-WSVLAEALDAVRHSVGLVWFLQHGVRKQADGSGSGSGT 411  
QY 405 POPNSKYKTSIMCRDLRQCGGPRGTNCTFAHSGEELEKYRLRNKKINATVTRTPFLNKKVG 464  
DB 412 TNSNPYKYSICRDLNVRVCPRGSGCTFAHSGEEVERARNR----- 455  
QY 455 VNNVTVTAGNVISVIGSTETTKVLPSTNGISNAENSVSLSIRSTSTLRALETVKV 524  
DB 456 ----- 455  
QY 525 GKVGVANGQNAAGPSADSVTENKIGSPKPTPSNVAAATSAGPSNVGTSLNSVP--OKSSPF 582  
DB 456 -----GKMKTPLA-----LOGPPAVGVGAIKKPLGEGEGP 487  
QY 583 LTRFVFPVPHSENIQYFODPRTQIPFVFPVQYPPPTPTVPAVAPCVPRFRSVNNV 642  
DB 583 LTRFVFPVPHSENIQYFODPRTQIPFVFPVQYPPPTPTVPAVAPCVPRFRSVNNV 642

```
Db 488 LGNMPMLPMPSP-MHYMGSPR-----GYLDP-----SLGLSP-----GGGL 522
QY 643 PESSLPPAS--MP--YADHYSTF---SPRDMNSSPYQPPPPQYGPVPPVPSGMYAPV 694
Db 523 PFSHISPIITRLVPSRYDSRFGGCTGR-----IPSPREYQANPVAP-----566
QY 695 YDSRIWRPMPYQORDIIRNSLPPM---DVMHSSVYQTSLRERNYSLDGYYSVACQPP- 750
Db 567 --TORANPNFNSVNSLNHGYMLPASGGDVFLA-----NPMQAYLAQQQHP 614
QY 751 -----SEPTTVPPLPREPCGHLKTSCEBQ-IRKPDQWQYHTQKAPLVS-STLPVATQS 803
Db 615 QHPQQQQPPSPKPNFSPRLSILPATADTSFFKCPNPSVSIIDLDRPVENVDAVPLFRS 674
QY 804 PTPPSPLFVDPRAADPSESVGTFKFEEDHLSHPWSCGTIGSCINAIIDSEPKDVIANSN 863
Db 675 -----NNNNNNNNNNNNNNNNNNHGSLLFPN-NTGKDSANFVRS---DSLDDDD 720
QY 864 AVLMDLDSGDVKKRRVHLFETQRTKEEDPIIPFSDGPIISKWGAISRSTRG 915
Db 721 ASTFDVPTGSSMLSIY-----GPICPK-----SSTTG 747

RESULT 11
Q7PTD3
ID Q7PTD3 PRELIMINARY; PRT; 398 AA.
AC Q7PTD3;
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DE ENSANGP0000011168 (Fragment).
GN Name=ENSANGG0000008679;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=PEST.
RC Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008807; EAA04070.2; -.
DR GO; GO:0003676; F: nucleic acid binding; IEA.
DR InterPro; IPR000571; Znf_CCH.
DR Pfam; PF00642; zf-CCH; 1.
FT NON_TER 1
FT NON_TER 398
SQ SEQUENCE 398 AA; 43739 MW; F1BCAFD08815A328 CRC64;

Query Match 20.7%; Score 1289.5; DB 2; Length 398;
Best Local Similarity 67.2%; Pred. No. 1.4e-56;
Matches 268; Conservative 38; Mismatches 68; Indels 25; Gaps 5;

QY 82 QSLKLNHKGKHYEYAKKVEDALYLKP-----LSGGKG-----VASLNQALS 127
Db 2 QSLSPENL---QYKSAKGCIIEALYLKPCPSAGSGSVAGSGGSGSPSGASLLS 58

QY 128 RPNQKLVTLVNCQVLEEGRVRAARSLGRTVTTELQHPNQQLSANLWAAVRAR 187
Db 59 RQKQKLVILVNCQLTEDEGRASLRARSLGRTVWELIHQNPQQLSTNLWAAVRAR 118

QY 188 GCQFLGPMQEEALKLVLLALEDGSALSRKVLVFWVQRLEPPQASKTSIGHVQVLLY 247
Db 119 GCQFLGPMQEEVLKLVLLALEDGSALSRKVLVFWVQRLEPPHFSQASKTSIGHVQVLLY 178

QY 248 RASCFKVKTKDESSLMQKEERSYEALRREDAQIVHITAMEAGLRISPEQWSLLYGD 307
Db 179 RASCFKVKSRGESSLMQKEERTYEALRREDAQIVQIATEAGRIAPDQWSLLYGD 238

QY 308 LAHKSQMOSIIDKQSPESPAKSQVELTIVLQRTGDPANLRLRPHLELLANIDPNDAV 367
```

```
Db 239 TAHKSHMQSINDKLTQTPQSFQSVQELIILQRTDPAKSLRVLQKHLAIDVNSML 298
QY 368 SPTVEQLENAVAVKTVVHGLVDFQNY-SRKGHETPQFPNPKYKTSKCRDLRQGGCP 426
Db 299 VPSMAECAALQAQKRVVAGLVDFVQHGNRKLQEAHLAHNRNYKISLCRDLNRRGTCP 358
QY 427 RGTNCTFAHSQEELEKYLKRNKK-----INATVTRTP 458
Db 359 RGPNCCTFAHSQEELEKYRKLKRSKSHGSSSSATASLRMVP 397

RESULT 12
Q9NNW9
ID Q9NNW9 PRELIMINARY; PRT; 198 AA.
AC Q9NNW9;
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DE 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ20713.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K.,
RA Hira M., Ohmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T.,
RA Shibahara T., Tanaka T., Nakamura Y., Isegaki T., Sugano S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK000720; BAA91340.1; -.
SQ SEQUENCE 198 AA; 22234 MW; E7778660861F1960 CRC64;

Query Match 14.3%; Score 889; DB 2; Length 198;
Best Local Similarity 98.9%; Pred. No. 5.3e-37;
Matches 173; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 867 MDLSDGDVKKRRVHLFETQRTKEEDPIIPFSDGPIISKWGAISRSTRGYHTTDPVQATA 926
Db 1 MDLSDGDVKKRRVHLFETQRTKEEDPIIPFSDGPIISKWGAISRSTRGYHTTDPVQATA 60

QY 927 SQGSATKPIVSVDYPPYNAVDSRWSSVGNATSSAHYVERDRFVTLDSGHRKHSSTGD 986
Db 61 SQGSATKPIVSVDYPPYNAVDSRWSSVGNATSSAHYVERDRFVTLDSGHRKHSSTGD 120

QY 987 LLSLELQQAQKSNLLOREANALAMQOKWNSLDEGRHLTLNLSKEIELRNGELQ 1041
Db 121 LLSLELQQAQKSNLLOREANALAMQOKWNSLDEGRHLTLNLSKEIELRNGEVK 175

RESULT 13
O45962
ID O45962 PRELIMINARY; PRT; 1014 AA.
AC O45962; Q93898;
DT 01-JUN-1998 (TEMBLrel. 06, Created)
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
DT 25-OCT-2004 (TEMBLrel. 28, Last annotation update)
DE Hypothetical protein M142.6.
GN ORFNames=M142.6;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
```



```
RC STRAIN=Bristol N2;
RA Steward C.A.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA McMurray A.A.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 273428; CAB16481.2; JOINED.
DR EMBL; 299276; CAB16481.2; -.
DR EMBL; 273428; CAA97810.2; -.
DR EMBL; 299276; CAA97810.2; JOINED.
DR F1R; T23764; T23764.
DR IntAct; O45962; -.
DR WormBase; WBGene00010923; M142.6.
DR WormPep; M142.6; CE34061.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0014567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; ZnF_ring.
DR PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR PROSITE; PS00589; ZF_RING_2; 1.
KW Hypothetical protein.
SQ
SEQUENCE 1014 AA; 112311 MW; 7226503B2DBE4008 CRC64;
Query Match 13.6%; Score 843; DB 2; Length 1014;
Best Local Similarity 30.6%; Pred. No. 1.1e-33;
Matches 238; Conservative 99; Mismatches 278; Indels 162; Gaps 23;
QY 4 QAAQTEFLSCPTCYNEFDENVHKKPISLGCSTHVCKTCLNKLHR--KACPFDDTAINTDI 61
DB 6 QGGQWQEVLCSSICNRHFE-TFLPVSLICGHVICKKAEKPNQTKPCDHDKTTHSP 64
QY 62 DVLVFNALLQLVGAQVPDQHSIKLNLGENKHGHEVAKKCVEDLALYKPLSGRGVAST 121
DB 65 SEYNNVALLSVI---FPRKQCTLSGAVSEAEKRVQDLSIQ-IAKFFREADSERG-GTV 119
QY 122 NSGALSRRPQKLVTLVNCQVEEGRVRAMRAASLGRRTVELILQHONQQQLSANLW 181
DB 120 SSREISRTLQKVLALLCYQREVQDGLTKLRCGRISERVMEIILTSQSNTHVSSQLW 179
QY 182 AAVRARGCQFLGPMQDEALKLVLALQEDGSALSRKVLVFLVVORLEPRFPQASKTSGH 241
DB 180 SAVRARGCQFLGPMQDDVLRLILMTLETGECIARKNLVNYVQTLASYPQVKTCVGH 239
QY 242 VVQLLYRASCFTKVRDESSLMQKEFRSYEARLRREHDAQIVHIAMEAGLRISPEQWS 301
DB 240 VVQLLYRASCNVLRKDGESSLMQKEFRFTVESLRREHDSQIVQIAPESGLRIGPDQWS 299
QY 302 SLLYGLDAKSHMQSIIDKLOSPESPAKSVOELTIVLQRTGDPAN-----LNLRP 352
DB 300 ALLYADQSHRSHMQSIIDKLOSKNSYQQGVVELRAL---AGSQTSLMLPAYRYFLTQVIP 356
QY 353 HLELLANIDPNDAVSPTEQLENAMVAVKTVVHGLVDFIONV---SRGHETPQPQNS 409
DB 357 CLEFFAGIEHEDTSNMIGDALHQIRLLK--LHCSQDDLKMKPEERGVLQAEVPGG 414
QY 410 KYKTSNCRDLRQGGCPRGNTCTTFAHSQBELEKYLRLNKKINATVTRFFLLNKVGNNTV 469
DB 415 -----MGGGPGSG-----GAAGRI----- 430
QY 470 TTAGNVISVIGSTETTKIVSTNGISNAENSVSQILSRSTDSTLRALETVKVKGKGA 529
DB 431 ----GGLHPLYSQIDETGRSISRTNPKNSHNSPQ----- 461
QY 530 NQONAAGPSADSVTENKIGSPKPTPVSNVAATSAGPSNVGTSLNSVPQKSSPLTRVPY 589
DB 462 -----TPPKQPKRYQMGPINRMG-----YSSDAPFFISHQOQ 497
QY 590 PPHSENIQFQDPRTOQIPFEVQYQPTGYPPPTVPAGVPCVPRFVRNNVPSSSLPP 649
DB 498 PPP---QPFNS--QHLPRFRGRGRGAPPPPPPPF-----MPMLIGYD-----MPG 539
```

```
QY 650 ASMPYADHYSTFSPDRMNSSP-----YQPPPPQYGVGPVPPSPGMYAPVYDSRRIRWPP 704
DB 540 APMQQATEVLT-ADQMGVNGTPQRVVMQSPHLPGGPVVWIPQOQMVPPOSMTVPVGGP 598
QY 705 MYQRDDIIRSNSLPMDVMHSSV-YQTSRLERYNSLDGYYSV---ACOPPSPPRTTV 757
DB 599 M-----GPMGPMTFIPVQVPPNTMTATSTPGSVIYPAASPPGQPPHTI 643

RESULT 14
Q96F27
ID Q96F27 PRELIMINARY; PRT; 177 AA.
AC Q96F27;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DE 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska J., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011688; AAH1688.2; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;
Query Match 12.6%; Score 780; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.3e-31;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 LNLKLRKACPPDOTAINTDIDVLPVNFALLQVGAQVPDQHSIKLNLGENKHGHEVAKK 101
DB 1 LNLKLRKACPPDOTAINTDIDVLPVNFALLQVGAQVPDQHSIKLNLGENKHGHEVAKK 60
QY 102 VEDLALYKPLSGGKGVASLNQSAISRPMQKRLVTLVNCQVEEGRVRAMRAASLGR 161
DB 61 VEDLALYKPLSGGKGVASLNQSAISRPMQKRLVTLVNCQVEEGRVRAMRAASLGR 120
QY 162 TVTEILQHONPQQLSANLWAAVRARGCQFLGP 194
DB 121 TVTEILQHONPQQLSANLWAAVRARGCQFLGP 153
```



## RESULT 15

```
Q8N9V1 PRELIMINARY; PRT; 522 AA.
AC Q8N9V1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ36182.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RC PubMed=14702039; DOI=10.1038/ngl1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kikawa E., Omura Y.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshioka W., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyaama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Yamazaki M., Watanabe T., Sugiyama A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45(2004).
DR EMBL; AK093501; BAC04186.1; --
SQ SEQUENCE 522 AA; 58253 MW; D4D774D3D00DD769 CRC64;

Query Match 9.8%; Score 611; DB 2; Length 522;
Best Local Similarity 34.1%; Pred. No. 1.6e-22;
Matches 187; Conservative 76; Mismatches 187; Indels 98; Gaps 27;

QY 618 YTPPTVPAGVAPCVFVRVSNVPESSLPASMEYADHYSTFSPDRMNSSPYQPPPP 677
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2 YTPPP-----QCVRFRV----PPSAPEAPYLDHYPPYL-QERVVNSQYGTQFQ 49

QY 678 QYGVPPVPSPGMYPVDSRIMWRPDMYQDDIIRNSLPPMDVMHSSV--YQTSLRER 735
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
50 Q-----YPII-----YPSHYDGRVYPAPSYTREIEPRESPI-PIETPPAAVPSYVPSRER 100

QY 736 YNSLDGYYSVACOP-----PS-----EPRTVPLPREPCGHLKTSCEQIRKPDQWAOYHT 787
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 YQIESYYPVAPHTQIRPSYLRPPYSLPPPPQP-----HPSLDELHRRRKEIMQAQLEE 156

QY 788 QKAPLVSSLTLPVATQPTPPSPFVSDFRADFSVSGTKFEDHLSHYSPWSCGTIGSC 847
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
157 RK---VISPPFPA-PSPTLP-PTFHPPEEFLDELDLVAG-KYKNDYQSYSPWSCDTIGSY 210

QY 848 INAIDSEPKDVIANSLVMDLDSGDVKKRVHLFETQRTKE--EDPIIPFSDGPIISKW 905
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
211 IGTKDAKPKDVVAAGSVEMNVESKGM--RDQRDLQRRAAETSDDDLIPFGDRPTVSRF 268
```

```
QY 906 GAISRSRTGYHTTDPVQATASQGSATKPISVSDYVPYVNAVDSRW--SSYG--NEATSS 961
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
269 GAISRTSKTIYQAGAPMQAMAPQAGTAKSINISDYSY--GTHGCGAGSPYSPHQNIPSQ 326
QY 962 AHYVERDRFIYDLSGHRK--HSSTGDLLSLELQQ-----AKSNLLQOR 1004
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
327 GHFSERERISMSEVASHGKPLPSAEREQLRLELQQLNHQISOQTOLRGPFAVNSRLVLQR 386
QY 1005 EANALAMQO-----KWNSLDEGRHILTLNL--LSKEIELRNGELQSDYTEDATDKPD 1054
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
387 EANTLAGSQSPPPPPPPKWPCKMISSEQLSLELHQVEREIGKTRLSME--NCSLDMK-- 443
QY 1055 RDIELELSALDTDE--PDQSEPIEELDIOLGISSQNDQLLNGVAVENGHPVQHQKEP 1112
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
444 -----SKLNTSKQAENGQPEPNKV-----PAEDLTLTFSDFVNGSALTQENISL 488
QY 1113 PKQKQSL 1120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
489 LSNKTSSL 496
```

Search completed: August 4, 2005, 01:40:03  
Job time : 197 secs

This Page Blank (uspt)